

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 11:16:20 ; Search time 26 Seconds

(without alignments)
332.692 Million cell updates/sec

Title: US-10-019-676-2

Perfect score: 451

Sequence: 1 STAGTAAKAIAGIKGKGE.....ATLQSSKNVYTSVNGQFTF 90

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	451	100.0	504 2 S33189	phase-1 flagellin
2	440	97.6	504 2 S33190	phase-1 flagellin
3	439	97.3	504 2 S33191	phase-1 flagellin
4	436	96.7	504 2 S33193	phase-1 flagellin
5	422	93.6	507 2 S33185	phase-1 flagellin
6	420	93.1	504 2 S33188	phase-1 flagellin
7	420	93.1	504 2 S33187	phase-1 flagellin
8	400	88.7	504 2 S33186	phase-1 flagellin
9	394	87.4	504 2 S33194	phase-1 flagellin
10	364	80.7	507 2 S33192	phase-1 flagellin
11	364	80.7	508 2 A53465	phase-1 flagellin
12	84	18.6	936 2 I40711	sapB protein - Cam
13	83.5	18.5	5189 2 B85547	probable RTX fam11
14	83.5	18.5	5291 2 F90696	hypothetical prote
15	82.5	18.3	572 2 G81277	flagellin Cj1338c
16	80.5	17.8	202 2 A59099	hypothetical prote
17	80.5	17.8	1109 2 A56143	surface-array prot
18	80.5	17.8	1461 2 B90966	hypothetical prote
19	80.5	17.8	1461 2 A85547	hypothetical prote
20	78.5	17.4	572 2 H81277	flagellin Cj1339c
21	78	17.3	729 2 T35028	probable glycosyl
22	77.5	17.2	394 2 A37853	flagellin, 40K - P
23	77	17.1	2554 2 A83528	extracellular seri
24	76	16.9	1477 2 B43855	high-molecular-wei
25	75.5	16.7	613 2 JC7827	X-Pro amonopeptida
26	75.5	16.7	857 2 B50009	cell wall surface
27	75.5	16.7	1161 2 C97881	conserved hypotnet
28	74.5	16.5	1072 2 G95851	probable hemolysin
29	74.5	16.5	1806 2 AF1717	probable peptidog1

30	74	16.4	1325	2	A64905	ydek protein - Esc
31	74	16.4	3029	2	S76109	hypothetical prote
32	73.5	16.3	912	2	D72644	hypothetical prote
33	73	16.2	474	2	S52444	flagellin A - legion
34	73	16.2	576	2	A39228	flagellin A - Camp
35	73	16.2	1343	2	B90893	hypothetical prote
36	73	16.2	1343	2	D85724	hypothetical prote
37	72.5	16.1	575	2	S41310	flagellin A - Camp
38	72.5	16.1	575	2	I40615	flagellin protein
39	72.5	16.1	635	2	S57714	capB protein - Clo
40	72.5	16.1	877	2	F90070	surface-factor B
41	72.5	16.1	893	2	A37284	surface-array prot
42	72	16.0	1741	2	S74910	hemolysin - Synch
43	72	16.0	2551	2	B98047	hypothetical prote
44	71.5	15.9	351	2	S60653	synape associated
45	71.5	15.9	377	2	B82112	flagellin Flad VC2

ALIGNMENTS

RESULT 1

S33189 phase-1 flagellin - Salmonella enteritidis (fragment)

C:Species: Salmonella enteritidis

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

C:Accession: S33189

R:Maaten, B.J.; Joys, T.M.

submitted to the EMBL Data Library, September 1992

A:Description: Molecular analysis of the flagellar antigen complex of Salmonella.

A:Reference number: S33185

A:Accession: S33189

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-504 <MAS>

A:Cross-references: EMBL:Z15068; NID:G296999; PIDN:CAA78777.1; PID:G297000

C:Superfamily: flagellin

Query Match 100.0%; Score 451; DB 2; Length 504;

Best Local Similarity 100.0%; Pred. No. 1.6e-34;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	1	STAGTAAKAIAGIKGKGEPTDYGKVTFTDTKTGDDNGKYSTTNGEKVTLIVAD	60
Db	252	STAGTAAKAIAGIKGKGEPTDYGKVTFTDTKTGDDNGKYSTTNGEKVTLIVAD	311
Oy	61	IATGATVNAATLQSSKNVYTSVNGQFTF	90
Db	312	IATGATVNAATLQSSKNVYTSVNGQFTF	341

RESULT 2

S33190 phase-1 flagellin - Salmonella montevideo (fragment)

C:Species: Salmonella montevideo

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

C:Accession: S33190

R:Maaten, B.J.; Joys, T.M.

submitted to the EMBL Data Library, September 1992

A:Description: Molecular analysis of the flagellar antigen complex of Salmonella.

A:Reference number: S33185

A:Accession: S33190

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-504 <MAS>

A:Cross-references: EMBL:Z15069; NID:G297002; PIDN:CAA78778.1; PID:G297003

C:Superfamily: flagellin

Oy	1	STAGTAAKAIAGIKGKGEPTDYGKVTFTDTKTGDDNGKYSTTNGEKVTLIVAD	60
Query Match	97.6%	Score 440; DB 2; Length 504;	
Best Local Similarity	97.8%	Pred. No. 1.7e-33;	
Matches	86; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	

DB 252 SAAGTDEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 3

S33191
phase-1 flagellin - Salmonella moscow (fragment)
C/Species: Salmonella moscow
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C/Accession: S33191
R/Masten, B.J.; Jovs, T.M.
submitted to the EMBL Data Library, September 1992
A/Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A/Reference number: S33185
A/Accession: S33191
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-504 <MAS>
A/Cross-references: EMBL:Z15086; NID:G297004; PIDN:CAA78794.1; PID:G297005
C/Superfamily: flagellin

Query Match 97.3%; Score 439; DB 2; Length 504;
Best Local Similarity 97.8%; Pred. No. 2.1e-33;
Matches 88; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 60
DB 252 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 4

S33193
phase-1 flagellin - Salmonella rostock (fragment)
C/Species: Salmonella rostock
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C/Accession: S33193
R/Masten, B.J.; Jovs, T.M.
submitted to the EMBL Data Library, September 1992
A/Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A/Reference number: S33185
A/Accession: S33193
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-504 <MAS>
A/Cross-references: EMBL:Z15071; NID:G297010; PIDN:CAA78780.1; PID:G297011
C/Superfamily: flagellin

Query Match 96.7%; Score 436; DB 2; Length 504;
Best Local Similarity 96.7%; Pred. No. 4e-33;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 60
DB 252 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 5

S33185
phase-1 flagellin - Salmonella berta (fragment)
C/Species: Salmonella berta
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

C/Accession: S33185
R/Masten, B.J.; Jovs, T.M.
submitted to the EMBL Data Library, September 1992
A/Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A/Reference number: S33185
A/Accession: S33185
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-507 <MAS>
A/Cross-references: EMBL:Z15064; NID:G297055; PIDN:CAA78773.1; PID:G297056
C/Superfamily: flagellin

Query Match 93.6%; Score 422; DB 2; Length 507;
Best Local Similarity 93.3%; Pred. No. 8.2e-32;
Matches 84; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 60
DB 252 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 314
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 315 IATGATDVNAATLQSSKNVYTSVVGQFTF 344

RESULT 6
S33188
phase-1 flagellin - Salmonella dublin (fragment)
C/Species: Salmonella dublin
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 12-Apr-1995

R/Masten, B.J.; Jovs, T.M.
submitted to the EMBL Data Library, September 1992
A/Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A/Reference number: S33185
A/Accession: S33188
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-504 <MAS>
A/Cross-references: EMBL:Z15067
C/Superfamily: flagellin

Query Match 93.1%; Score 420; DB 2; Length 504;
Best Local Similarity 93.3%; Pred. No. 1.3e-31;
Matches 84; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 60
DB 252 STAGTAEAKAIAIGAKGEGDTPDKGVTFTIDTGTGGDNGKYSTTINGEKVTLTVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 7

S33187
phase-1 flagellin - Salmonella derby (fragment)
C/Species: Salmonella derby
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Mar-1998
C/Accession: S33187
R/Masten, B.J.; Jovs, T.M.
submitted to the EMBL Data Library, September 1992
A/Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A/Reference number: S33185
A/Accession: S33187
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-504 <MAS>
A/Cross-references: EMBL:Z15066
C/Superfamily: flagellin

Query Match 93.1%; Score 420; DB 2; Length 504;

A:Molecule type: DNA
A:Residues: 1-936 <RES>
A:Cross-references: EMBL:U25133; NID:G801999; PIDN:AAA79683.1; PID:G802000
C:Genetics:
A:Gene: sapB

Query Match 18.6%; Score 84; DB 2; Length 936;
Best Local Similarity 30.8%; Pred. No. 5;
Matches 28; Conservative 9; Mismatches 36; Indels 18; Gaps 5;

QY 5 TAEAKAIA---GAIKGKEGDPFDYKGVTFITDTKTGDD-----GNGKYS-TTIN-- 50
DB 778 TPDASAITLGANNTIKGGSGADSIYKGNIVDVLAGDDDTTLTKKGRKEDITTVNPF 837

QY 51 --GEKVTLVADIATATVNAATLQSSKNV 79
DB 838 NAGDKIDIT--DAKNGTFENKTIWNNDNL 866

RESULT 13

B85547
probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, substrain C);Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85547
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.C.; Mayhew Miller, L.F.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 403, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85547
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5188 <STO>
A:Cross-references: GB:AE005174; NID:G12513368; PIDN:AAG54838.1; GSPDB:GN00145; UMGPI:Z06
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0615

Query Match 18.5%; Score 83.5; DB 2; Length 5188;
Best Local Similarity 27.8%; Pred. No. 35;
Matches 25; Conservative 12; Mismatches 28; Indels 25; Gaps 2;

QY 1 STAGTAEAKAIAIGAIKGEKDPFDYKGVTFITDTKTGDDGNGKYSTTINGEKVTLTVAD 60
DB 653 NAAESGNAGTISGQVTGAAGSDT-----VTVLGNTYATVQS 691

QY 61 IATGATVNAATLQSSKN---VYTSVNG 86
DB 692 NLSWSVDVPAADIQALGNGDLTVNASVTNG 721

RESULT 14

F90696
hypothetical protein ECG0542 [imported] - Escherichia coli (strain O157:H7, substrain R1 C);Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90696
R:Hayashi, T.; Makino, K.; Omishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gatawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingawa, H. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156211; PMID:11258796
A:Accession: F90696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5291 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833965.1; PID:G13360000; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECG0542

Query Match 18.5%; Score 83.5; DB 2; Length 5291;

Best Local Similarity 27.8%; Pred. No. 36;
Matches 25; Conservative 12; Mismatches 28; Indels 25; Gaps 2;

QY 1 STAGTAEAKAIAIGAIKGEKDPFDYKGVTFITDTKTGDDGNGKYSTTINGEKVTLTVAD 60
DB 653 NAAESGNAGTISGQVTGAAGSDT-----VTVLGNTYATVQS 691

QY 61 IATGATVNAATLQSSKN---VYTSVNG 86
DB 692 NLSWSVDVPAADIQALGNGDLTVNASVTNG 721

RESULT 15

G81277
flagellin Cj1338c [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: G81277
R:Parhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chill C.M.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barr Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: G81277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:G6968723; PIDN:CAB73765.1; PID:G6968
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: flab; Cj1338c
C:Superfamily: flagellin

Query Match 18.3%; Score 82.5; DB 2; Length 572;
Best Local Similarity 34.2%; Pred. No. 4.1;
Matches 25; Conservative 7; Mismatches 36; Indels 5; Gaps 2;

QY 5 TAEAKAIAIGAIKGEKDPFDYKGVTFITDTKTGDDGNGKYSTTINGEKVTLTVADIAG 64
DB 235 TVEIRGIA-AVRAAGTSDTFAINGVTIQVAYEDDGALVAAINSVADITGV---EA 289

QY 65 ATDVNAATLQSSK 77
DB 230 STDANGQLLNSR 302

Search completed: January 12, 2004, 11:18:23
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 11:16:21 ; Search time 54 Seconds

(without alignments)
430.088 Million cell updates/sec

Title: US-10-019-676-2

Perfect score: 451

Sequence: 1 STAGTAAKAIAGIKGKKE.....ATLQSSKNVYTSVNGQFTF 90

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451	100.0	494	2	Q54210
2	447	99.1	505	2	Q54863
3	447	99.1	505	2	Q54864
4	446	98.9	465	2	Q53967
5	446	98.9	493	2	Q53998
6	446	98.9	504	2	Q53989
7	446	98.9	504	2	Q54329
8	446	98.9	504	2	Q53822
9	446	98.9	505	2	Q53970
10	445	98.7	505	2	Q53781
11	434	96.2	504	2	Q53993
12	422	93.6	504	2	Q53583
13	417	92.5	505	2	Q53990
14	406	90.0	504	2	Q54489
15	405	89.8	504	2	Q53991
16	405	89.8	504	2	Q53996

17	401	88.9	503	2	Q54414	Q54414 salmonella
18	397	88.0	505	2	Q53992	Q53992 salmonella
19	375.5	83.3	504	2	Q53994	Q53994 salmonella
20	373	82.7	508	2	Q53995	Q53995 salmonella
21	372	82.5	507	2	Q53821	Q53821 salmonella
22	367	81.4	507	2	Q54415	Q54415 salmonella
23	353	78.3	508	2	Q54515	Q54515 salmonella
24	224	49.7	52	2	Q9R2V0	Q9R2V0 salmonella
25	217	48.1	45	2	Q9R405	Q9R405 salmonella
26	213	47.2	45	2	Q9R406	Q9R406 salmonella
27	127	28.2	449	2	Q8GG11	Q8GG11 escherichia
28	127	28.2	456	2	Q9R308	Q9R308 escherichia
29	102.5	22.7	454	2	Q8GG12	Q8GG12 escherichia
30	102	22.6	455	2	Q8GGH8	Q8GGH8 escherichia
31	94.5	21.0	487	2	Q9J8E4	Q9J8E4 escherichia
32	92.5	20.5	448	2	Q8GGH9	Q8GGH9 escherichia
33	90.5	20.1	452	2	Q8GG10	Q8GG10 escherichia
34	85.5	19.0	2353	2	P71401	P71401 haemophilus
35	84	18.6	936	2	Q46037	Q46037 campylobact
36	83.5	18.5	1557	2	Q9RNT2	Q9RNT2 haemophilus
37	83.5	18.5	5188	16	Q8X4H5	Q8X4H5 escherichia
38	83.5	18.5	5291	16	Q8X2T1	Q8X2T1 escherichia
39	83	18.4	1758	5	Q811L6	Q811L6 plasmodium
40	83	18.4	1758	5	Q811L5	Q811L5 plasmodium
41	82.5	18.3	572	2	Q9R9E1	Q9R9E1 campylobact
42	82.5	18.3	572	2	Q9R9E4	Q9R9E4 campylobact
43	81	18.0	2035	2	Q9XC04	Q9XC04 salmonella
44	81	18.0	2039	16	Q8ZM57	Q8ZM57 salmonella
45	80.5	17.8	201	2	Q8KXN7	Q8KXN7 bacillus an

ALIGNMENTS

ID	Q54210	PRELIMINARY;	PRT;	494 AA.
AC	Q54210;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Phase-1 flagellin.			
GN	FlaC1.			
OS	Salmonella gallinarum			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Salmonella.			
OX	NCBI_TaxID=594;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEMLINE=3156031; PubMed=8429538;			
RA	Li J., Smith N.H., Nelson K., Crichton P.B., Old D.C., Whitlam T.S.,			
RA	Selander R.K.;			
RT	"Evolutionary origin and radiation of the avian-adapted non-motile			
RT	Salmonellae."			
RL	J. Med. Microbiol. 38:129-139 (1993).			
DR	EMBL; M84975; AAA27085.1; -			
DR	InterPro; IPR001492; flagellin_N.			
DR	InterPro; IPR001029; flagellin_C.			
DR	Pfam; PF00700; flagellin_C; 1.			
DR	Pfam; PF00669; flagellin_N; 1.			
DR	PRINTS; PR00207; FLAGELLIN.			
DR	ProDom; PD000316; flagellin_C; 2.			
SO	SEQUENCE 494 AA; 51732 MW; 1DD4879A854AD73 CRC64;			
Query Match		100.0%;	Score 451;	DB 2;
Best Local Similarity		100.0%;	Pred. No. 7,8e-30;	Length 494;
Matches 90;		Conservative 0;	Mismatches 0;	Indels 0;
Gaps 0;				
DB	1 STAGTAAKAIAGIKGKGDTPDYKGVPTTITTKGDDNGKGVSTTINSEKTLIVAD 60			
DB	253 STAGTAAKAIAGIKGKGDTPDYKGVPTTITTKGDDNGKGVSTTINSEKTLIVAD 312			
QY	61 IATGATVNAATLQSSKNVYTSVNGQFTF 90			

```

Db      313  IATGATDVNAATLQSSKNVYTSVNGQFTF 342
|||||
RESULT 2
Q54863  PRELIMINARY; PRT; 505 AA.
AC      Q54863;
DT      01-NOV-1996 (TEMBLrel. 01, Created)
DT      01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin.
GN      FlIC.
OS      Salmonella pullorum.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=605;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93156031; PubMed=8429538;
RA      Li J., Smith N.H., Nelson K., Crichton P.B., Old D.C., Whittam T.S.,
RA      Selander R.K.,
RT      "Evolutionary origin and radiation of the avian-adapted non-motile
RT      Salmonellae."
RL      J. Med. Microbiol. 38:129-139(1993).
DR      EMBL; M84977; AAA27087.1; -
DR      InterPro: IPR001492; Flagellin_N.
DR      InterPro: IPR001029; Flagellin_C.
DR      Pfam; PF00700; Flagellin_C; 1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; FLAGELLIN.
DR      ProDom; PD000316; Flagellin_C; 2.
SQ      SEQUENCE 505 AA; 52912 MW; 3C7567432598D81 CRC64;

Query Match      99.1%; Score 447; DB 2; Length 505;
Best Local Similarity 98.9%; Pred. No. 1.7e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 60
Db      253 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 312
|||||

QY      61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db      313 IATGATDVNAATLQSSKNVYTSVNGKFTF 342
|||||

RESULT 3
Q54864  PRELIMINARY; PRT; 505 AA.
AC      Q54864;
DT      01-NOV-1996 (TEMBLrel. 01, Created)
DT      01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin.
GN      FlIC.
OS      Salmonella pullorum.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=605;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93156031; PubMed=8429538;
RA      Li J., Smith N.H., Nelson K., Crichton P.B., Old D.C., Whittam T.S.,
RA      Selander R.K.,
RT      "Evolutionary origin and radiation of the avian-adapted non-motile
RT      Salmonellae."
RL      J. Med. Microbiol. 38:129-139(1993).
DR      EMBL; M84978; AAA27088.1; -
DR      InterPro: IPR001492; Flagellin_N.
DR      InterPro: IPR001029; Flagellin_C.
DR      Pfam; PF00700; Flagellin_C; 1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; FLAGELLIN_N; 1.
DR      ProDom; PD000316; Flagellin_N; 1.
SQ      SEQUENCE 505 AA; 52912 MW; 3C7567432598D81 CRC64;

```

```

DR      PRINTS; PR00207; FLAGELLIN.
DR      ProDom; PD000316; Flagellin_C; 2.
SQ      SEQUENCE 505 AA; 52942 MW; 3C75770533499C80 CRC64;

Query Match      99.1%; Score 447; DB 2; Length 505;
Best Local Similarity 98.9%; Pred. No. 1.7e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 60
Db      253 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 312
|||||

QY      61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db      313 IATGATDVNAATLQSSKNVYTSVNGKFTF 342
|||||

RESULT 4
Q53967  PRELIMINARY; PRT; 465 AA.
AC      Q53967;
DT      01-NOV-1996 (TEMBLrel. 01, Created)
DT      01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin (Fragment).
GN      FlIC.
OS      Salmonella enterica subsp. enterica serovar Enteritidis var. danysz.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=29476;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=ATCC 49216;
RA      Masten B.J., Jorys T.M.;
RA      Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RL      EMBL; U05298; AAA53493.1; -
DR      InterPro: IPR001492; Flagellin_N.
DR      InterPro: IPR001029; Flagellin_C.
DR      Pfam; PF00700; Flagellin_C; 1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; FLAGELLIN.
DR      ProDom; PD000316; Flagellin_C; 2.
FT      NON TER
SQ      SEQUENCE 465 AA; 48938 MW; EE7A26A8FA429A1 CRC64;

Query Match      98.9%; Score 446; DB 2; Length 465;
Best Local Similarity 98.9%; Pred. No. 1.9e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 60
Db      252 STAGTAEAKAIAAGAIKGGKEGDPYKGVTFITIDTKTGDDGNGKYSTTINGEKVTLTVAD 311
|||||

QY      61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db      312 IATGATDVNAATLQSSKNVYTSVNGQFTF 341
|||||

RESULT 5
Q53998  PRELIMINARY; PRT; 493 AA.
AC      Q53998;
DT      01-NOV-1996 (TEMBLrel. 01, Created)
DT      01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin (Fragment).
GN      FlIC.
OS      Salmonella enteritidis.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=592;
RN      [1]
RP      SEQUENCE FROM N.A.

```

```

RC STRAIN=857;
RX MEDLINE=95189740; PubMed=7533759;
RA van Asten A.J.A.M., Zwaagstra K.A., Baay M.F., Kuusters J.G.,
RA Huis in 't Veld J.H.J., van der Zeijst B.A.M.,
RT "Identification of the domain which determines the Gm-serotype of the
RT flagellin of Salmonella enteritidis.";
RL J. Bacteriol. 177:1610-1613 (1995).
DR EMBL; U12963; AAA64387.1;
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_N; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 2.
FT NON_TER 1
SQ SEQUENCE 493 AA; 51794 MW; 4387EC9DEFBD400B CRC64;

Query Match 98.9%; Score 446; DB 2; Length 493;
Best Local Similarity 98.9%; Pred. No. 2e-29;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 60
Db 241 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 300

QY 61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db 301 IATGATDVNAATLQSSKNVYTSVNGQFTF 330

RESULT 6
O53989 PRELIMINARY; PRT; 504 AA.
AC O53989;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phase-1 flagellin (Fragment).
GN FlgC.
OS Salmonella enterica subsp. enterica serovar Basen.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=29477;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49219;
RA Maaten B.J., Joys T.M.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U05299; AAA53494.1;
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_N; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 2.
FT NON_TER 1
SQ SEQUENCE 504 AA; 52849 MW; AEA37AAE46E3EEF CRC64;

Query Match 98.9%; Score 446; DB 2; Length 504;
Best Local Similarity 98.9%; Pred. No. 2.1e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 60
Db 252 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 311

QY 61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db 312 IATGATDVNAATLQSSKNVYTSVNGQFTF 341

RESULT 7
O54329

```

```

ID O54329 PRELIMINARY; PRT; 504 AA.
AC O54329;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phase-1 flagellin (Fragment).
GN FlgC.
OS Salmonella enterica subsp. enterica serovar Enteritidis var. Jena.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=29478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49221;
RA Maaten B.J., Joys T.M.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U05300; AAA53495.1;
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_N; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 2.
FT NON_TER 1
SQ SEQUENCE 504 AA; 52849 MW; AEA37AAE46E3EEF CRC64;

Query Match 98.9%; Score 446; DB 2; Length 504;
Best Local Similarity 98.9%; Pred. No. 2.1e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 60
Db 252 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 311

QY 61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90
Db 312 IATGATDVNAATLQSSKNVYTSVNGQFTF 341

RESULT 8
O53822 PRELIMINARY; PRT; 504 AA.
AC O53822;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Phase-1 flagellin (Fragment).
GN FlgC.
OS Salmonella enterica subsp. enterica serovar Enteritidis var. Chaco.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=29475;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49214;
RA Maaten B.J., Joys T.M.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U05297; AAA53492.1;
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_N; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 2.
FT NON_TER 1
SQ SEQUENCE 504 AA; 52849 MW; AEA37AAE46E3EEF CRC64;

Query Match 98.9%; Score 446; DB 2; Length 504;
Best Local Similarity 98.9%; Pred. No. 2.1e-29;
Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIAAGKGGKEDTDFYKGVFTTIDTKGDDGNGKVSSTINGEKVTLTVAD 60

```

```

Db      252 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 311
Qy      61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db      312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 9
ID      053970      PRELIMINARY;      PRT;      505 AA.
AC      Q53970;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin.
GN      FljC1.
OS      Salmonella dublin.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxId=98360;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92276338; PubMed=1592813;
RA      Selander R.K., Smith N.H., Li J., Beltran P., Ferris K.E.,
RA      Kopecko D.J., Rubin F.A.;
RT      "Molecular evolutionary genetics of the cattle-adapted serovar
RT      Salmonella dublin.";
SQ      PRODOM: PD000316; FLAGELLIN_C; 2.
SQ      SEQUENCE 505 AA; 52951 MW; 01A5C6778B275F5C CRC64;

Query Match
Best Local Similarity 98.9%; Score 446; DB 2; Length 505;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 60
Db      253 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 312
Qy      61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db      313 IATGATDVNAATLQSSKNVYTSVVGQFTF 342

RESULT 10
ID      057381      PRELIMINARY;      PRT;      505 AA.
AC      Q57381;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Phase-1 flagellin.
GN      FljC1.
OS      Salmonella enteritidis, and
OS      Salmonella enterica subsp. enterica serovar Othmarachen.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxId=592; 34041;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92276338; PubMed=1592813;
RA      Selander R.K., Smith N.H., Li J., Beltran P., Ferris K.E.,
RA      Kopecko D.J., Rubin F.A.;
RT      "Molecular evolutionary genetics of the cattle-adapted serovar
RT      Salmonella dublin.";
SQ      PRODOM: PD000316; FLAGELLIN_C; 2.
SQ      SEQUENCE 508 AA; 53332 MW; 3A6A23F526159B8C CRC64;

Query Match
Best Local Similarity 96.2%; Score 434; DB 2; Length 508;
Matches 86; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 60
Db      256 SAVGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 315
Qy      61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db      316 IATGATDVNAATLQSSKNVYTSVVGQFTF 345

RESULT 12
ID      053583

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=S5334;
RX      MEDLINE=94195780; PubMed=8146152;
RA      Li J., Nelson K., McWhorter A.C., Whittam T.S., Selander R.K.;
RA      "Recombinational basis of serovar diversity in Salmonella enterica.";
RT      Proc. Natl. Acad. Sci. U.S.A. 91:2552-2556(1994).
DR      EMBL; M84980; AAA27092.1; -
DR      EMBL; U06455; AAA17869.1; -
DR      InterPro; IPR001492; Flagellin_N.
DR      InterPro; IPR001029; Flagellin_C.
DR      Pfam; PF00700; Flagellin_N; 1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; FLAGELLIN.
DR      PRODOM: PD000316; FLAGELLIN_C; 2.
SQ      SEQUENCE 505 AA; 53011 MW; F99D43CEB3E74D8C CRC64;

Query Match
Best Local Similarity 98.7%; Score 445; DB 2; Length 505;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 60
Db      253 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 312
Qy      61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db      313 IATGATDVNAATLQSSKNVYTSVVGQFTF 342

RESULT 11
ID      053993      PRELIMINARY;      PRT;      508 AA.
AC      Q53993;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Phase 1 flagellin.
GN      FljC.
OS      Salmonella enterica.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxId=28901;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=S5335;
RX      MEDLINE=94195780; PubMed=8146152;
RA      Li J., Nelson K., McWhorter A.C., Whittam T.S., Selander R.K.;
RA      "Recombinational basis of serovar diversity in Salmonella enterica.";
RT      Proc. Natl. Acad. Sci. U.S.A. 91:2552-2556(1994).
DR      EMBL; U06204; AAA17863.1; -
DR      InterPro; IPR001492; Flagellin_N.
DR      InterPro; IPR001029; Flagellin_C.
DR      Pfam; PF00700; Flagellin_N; 1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; FLAGELLIN.
DR      PRODOM: PD000316; FLAGELLIN_C; 2.
SQ      SEQUENCE 508 AA; 53332 MW; 3A6A23F526159B8C CRC64;

Query Match
Best Local Similarity 95.6%; Score 434; DB 2; Length 508;
Matches 86; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 STAGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 60
Db      256 SAVGTAEKAIAGAIKGGKEDTDPYKGVTFITIDTKTGDNGKYSTTINGEKVLTIVAD 315
Qy      61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db      316 IATGATDVNAATLQSSKNVYTSVVGQFTF 345

```


DR Pfam; PF00700; Flagellin C; 1.
DR Pfam; PF00669; Flagellin N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR PRODOM; PD000316; Flagellin C; 2.
SQ SEQUENCE 504 AA; 52932 MW; 000ACF05AC678FF CRC64;

Query Match 89.8%; Score 405; DB 2; Length 504;
Best Local Similarity 90.0%; Pred. No. 5.5e-26;
Matches 81; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIAGAIKGGKGGTFFDYKGYTFITIDTKGDDGNGKYSTTINGEKVTLTVAD 60
Db 252 SSAGTAESKAIATAIKGGKGGTFFDYKGYTFITIDTKGDDGNGKYSTTINGEKVTLTVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
Db 312 ITTGADVNAATLQSSKNVYTSVVGQFTF 341

Search completed: January 12, 2004, 11:19:28
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 11:16:17 ; Search time 18 seconds

(Without alignments)
235.134 Million cell updates/sec

Title: US-10-019-676-2

Sequence: 1 STAGTAFAKAIAGAIKGRKE.....ATLQSSKNVYTSVNGQFTF 90

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	451	100.0	504	1	FLIC_SALEN
2	440	97.6	504	1	FLIC_SALDU
3	440	97.6	504	1	FLIC_SALMO
4	439	97.3	504	1	FLIC_SALMO
5	436	96.7	504	1	FLIC_SALMO
6	435	96.5	504	1	FLIC_SALMO
7	422	93.6	507	1	FLIC_SALMO
8	420	93.1	504	1	FLIC_SALMO
9	400	88.7	504	1	FLIC_SALMO
10	394	87.4	504	1	FLIC_SALMO
11	364	80.7	507	1	FLIC_SALMO
12	82.5	18.3	469	1	ALSA_CANAL
13	82.5	18.3	571	1	FLAB_CAMU
14	82.5	18.3	571	1	FLAB_CAMU
15	78.5	17.4	571	1	FLAB_CAMU
16	77.5	17.2	333	1	FLAB_CAMU
17	75	16.6	334	1	FLAB_CAMU
18	73	16.4	1335	1	YDEK_BCOI
19	73	16.2	474	1	FLAB_CAMU
20	73	16.2	575	1	FLAB_CAMU
21	72.5	16.1	574	1	FLAB_CAMU
22	71.5	15.9	376	1	FLAB_CAMU
23	70.5	15.6	1902	1	P1P_LACIC
24	70.5	15.6	1902	1	P1P_LACIC
25	70.5	15.6	1902	1	P1P_LACIC
26	69.5	15.4	1902	1	P1P_LACIC
27	69.5	15.4	572	1	FLAB_CAMU
28	69	15.3	528	1	FLAB_CAMU
29	69	15.3	528	1	FLAB_CAMU
30	68.5	15.2	775	1	YTXI_XENIA
31	68.5	15.2	4590	1	P14380 xenopus lae
32	68	15.1	209	1	PYRE_STRAS
33	68	15.1	299	1	PRTX_LACIC

ALIGNMENTS

34	68	15.1	340	1	THPS_SULAC
35	68	15.1	367	1	TRMU_MYCTU
36	67.5	15.0	445	1	FLUD_VIBPA
37	67.5	15.0	547	1	NLTP_CHICK
38	67.5	15.0	642	1	FLUD_CAMU
39	67	14.9	592	1	UUP_BUCAP
40	67	14.9	1296	1	ASAI_ENTFA
41	67	14.9	2021	1	OMPA_RICCN
42	67	14.9	2358	1	YEBI_BCOI
43	66.5	14.7	572	1	FLAB_CAMU
44	66.5	14.7	643	1	HS76_HUMAN
45	66.5	14.7	2647	1	FLNA_HUMAN

RESULT 1
FLIC_SALEN
ID FLIC_SALEN STANDARD; PRT; 504 AA.
AC 006972;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC.
OS Salmonella enteritidis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13076;
RX MEDLINE=93374829; PubMed=7690024;
RA Maaten B.J., Joys T.M.;
RT "Molecular analyses of the Salmonella g. . . flagellar antigen complex."
RL J. Bacteriol. 175:5359-5365(1993).
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLIN, TERMED PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: Z15068; CAA7877.1; -.
CC PIR: S3189; S33189.
DR InterPro: IPR001029; Flagellin C.
DR InterPro: IPR001492; Flagellin N.
DR Pfam: PR00700; Flagellin_C_1.
DR Pfam: PR00669; Flagellin_N_1.
DR PRINTS: PR00207; FLAGELLIN.
KW Flagella.
FT INIT MET.
SQ SEQUENCE 504 AA; 52850 MW; 4040770EE2E2B985 CRC64;
BY SIMILARITY.
Query Match 100.0%; Score 451; DB 1; Length 504;
Best local Similarity 100.0%; Pred. No. 2.7e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STAGTAFAKAIAGAIKGRKEDTFFYKGYFTIDFTKGDGNGKYSTINGEKVTLVAD 60
DB 252 STAGTAFAKAIAGAIKGRKESDTPFYKGYFTIDFTKGDGNGKYSTINGEKVTLVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVNGQFTF 90

DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 2

ID FLIC_SALDU STANDARD; PRT; 504 AA.
AC 006971;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC OR FLICI.
OS Salmonella dublin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=98360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92276338; PubMed=1592813;
RA Selandar R.K., Smith N.H., Li J., Beltan P., Ferris K.E.,
RT Kopecsko D.J., Rubin F.A.,
RT "Molecular evolutionary genetics of the cattle-adapted serovar
Salmonella dublin."
RT J. Bacteriol. 174:3587-3592(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15480;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.J., Joys T.M.,
RT "Molecular analyses of the Salmonella g. . . flagellar antigen
complex."
RT J. Bacteriol. 175:5359-5365(1993).
RN [1]
RP FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M84973; AAA27081.1; -
DR EMBL; Z15067; CAA78776.1; -
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C.1.
DR Pfam: PF00669; Flagellin_N.1.
DR PRINTS; PR00207; FLAGELLIN.
DR PRINTS; PR00207; FLAGELLIN.
KM Flagella.
FT INIT MET
SQ SEQUENCE 504 AA; 52860 MW; 34AE0DB50330AFC CRC64;
BY SIMILARITY.

Query Match 97.6%; Score 440; DB 1; Length 504;
Best Local Similarity 97.8%; Pred. No. 2.7e-33;
Matches 88; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAGTAEKKAIAAGAKGKSGEDTDFYGVFTIDTKGDDGNGVSTTINSEKVTLYVAD 60
DB 252 STAGTAEKKAIAAGAKGKSGEDTDFYGVFTIDTKGDDGNGVSTTINSEKVTLYVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 3

ID FLIC_SALMO STANDARD; PRT; 504 AA.
AC 006973;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC.
OS Salmonella montevideo.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=115981;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 8387;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.J., Joys T.M.,
RT "Molecular analyses of the Salmonella g. . . flagellar antigen
complex."
RT J. Bacteriol. 175:5359-5365(1993).
RN [1]
RP FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; Z15069; CAA78778.1; -
DR EMBL; S33190; S33190.
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C.1.
DR Pfam: PF00669; Flagellin_N.1.
DR PRINTS; PR00207; FLAGELLIN.
DR PRINTS; PR00207; FLAGELLIN.
KM Flagella.
FT INIT MET
SQ SEQUENCE 504 AA; 52862 MW; 5C090577E21E6A67 CRC64;
BY SIMILARITY.

Query Match 97.6%; Score 440; DB 1; Length 504;
Best Local Similarity 97.8%; Pred. No. 2.7e-33;
Matches 88; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 STAGTAEKKAIAAGAKGKSGEDTDFYGVFTIDTKGDDGNGVSTTINSEKVTLYVAD 60
DB 252 SAAGTDEKKAIAAGAKGKSGEDTDFYGVFTIDTKGDDGNGVSTTINSEKVTLYVAD 311
QY 61 IATGATDVNAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVNAATLQSSKNVYTSVVGQFTF 341

RESULT 4

ID FLIC_SALMC STANDARD; PRT; 504 AA.
AC 006981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-B flagellin).
GN FLIC.
OS Salmonella moscow.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=28146;
RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=CDC;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.J., Jøys T.M.;
RT "Molecular analyses of the Salmonella g. . . flagellar antigen
complex.";
RL J. Bacteriol. 175:5359-5365(1993).
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Z15086; CAA78794.1; -
DR PIR: S33191; S33191.
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C_1.
DR Pfam: PF00669; Flagellin_N_1.
DR PRINTS: PR00207; FLAGELLIN.
KW Flagella.
FT INIT MET
SQ SEQUENCE 504 AA; 52791 MW; FCBA2180AF11A0 CRC64;

Query Match
Best Local Similarity 97.3%; Score 439; DB 1; Length 504;
Matches 88; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIGAKIGKESDTPDKGYTFITDTKTGGDNGKYSTTNGEKVTLTVAD 60
DB 252 STAGTAARAKAIGAKIGKESDTPDKGYTFITDTKTGGDNGKYSTTNGEKVTLTVAD 311
QY 61 IATGATDVAAATLQSSKNVYTSVNGQFTF 90
DB 312 IATGATDVAAATLQSSKNVYTSVNGQFTF 341

RESULT 5
FLIC_SALMO STANDARD; PRT; 504 AA.
ID FLIC_SALMO
AC Q06982;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC.
OS Salmonella rostock.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=28149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.J., Jøys T.M.;
RT "Molecular analyses of the Salmonella g. . . flagellar antigen
complex.";
RL J. Bacteriol. 175:5359-5365(1993).
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Z15071; CAA78780.1; -
DR PIR: S33193; S33193.
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C_1.
DR Pfam: PF00669; Flagellin_N_1.
DR PRINTS: PR00207; FLAGELLIN.
KW Flagella.
FT INIT MET
SQ SEQUENCE 504 AA; 52846 MW; DFOC32476B93DF5F CRC64;

Query Match
Best Local Similarity 96.7%; Score 436; DB 1; Length 504;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAARAKAIGAKIGKESDTPDKGYTFITDTKTGGDNGKYSTTNGEKVTLTVAD 60
DB 252 STAGTAARAKAIGAKIGKESDTPDKGYTFITDTKTGGDNGKYSTTNGEKVTLTVAD 311
QY 61 IATGATDVAAATLQSSKNVYTSVNGQFTF 90
DB 312 IATGATDVAAATLQSSKNVYTSVNGQFTF 341

RESULT 6
FLIC_SALINA STANDARD; PRT; 504 AA.
ID FLIC_SALINA
AC Q52959;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Phase-1 flagellin.
GN FLIC.
OS Salmonella naestved.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=71517;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L-5;
RA Sekizaki T., Sato Y., Osaki M., Mitsumori M.;
RT "Detection and typing of Salmonella fljC gene by PCR-RFLP for
RT diagnosis of bovine salmonellosis.";
RT Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D78639; BAA24529.1; -
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C_1.
DR Pfam: PF00669; Flagellin_N_1.

```

DR PRINTS; PRO0207; FLAGELLIN.
KW Flagella.
SQ INIT MET 0 0 BY SIMILARITY.
SEQUENCE 504 AA; 52803 MW; 42A0FCEBFJ1152A CRC64;

Query Match
Best Local Similarity 96.5%; Score 435; DB 1; Length 504;
Matches 87; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAERAKAIAIGKGGEDTDPYKGVTFITDITKTGDDNGKYSTTINGEKVTLTVAD 60
DB 252 STAGTAERAKAIAIGKGGEDTDPYKGVTFITDITKTGDDNGKYSTTINGEKVTLTVAD 311

QY 61 IATGATDVAAATLQSSKNVYTSVVGQFTF 90
DB 312 IATGATDVAAATLQSSKNVYTSVVGQFTF 341

RESULT 7
FLIC_SALBE STANDARD; PRT; 507 AA.
AC 006968;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-I flagellin).
GN FLIC.
OS Salmoneilla berta.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneilla.
OX NCBI_TaxID=28144;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-S5321;
RX MEDLINE=94195780; PubMed=8146152; Whitlam T.S., Selander R.K.,
RA Li J., Nelson K., McWhorter A.C., Whitlam T.S., Selander R.K.,
RT "Recombinational basis of serovar diversity in Salmoneilla enterica.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:2552-2556(1994).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 8392;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.U., Jovs T.M.;
RT "Molecular analyses of the Salmoneilla g. . . flagellar antigen
complex.";
RL J. Bacteriol. 175:5359-5365(1993).
RN (3)
RP FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SARMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U06227; AAA17868.1; -
CC EMBL; Z15064; CAA78773.1; -
CC InterPro; IPR001029; Flagellin_C.
CC InterPro; IPR001492; Flagellin_N.
CC Pfam; PF00700; Flagellin_C_1.
CC Pfam; PF00669; Flagellin_N_1.
DR PRINTS; PRO0207; FLAGELLIN.
KW Flagella.
FT INIT MET 0 0 BY SIMILARITY.
SEQUENCE 507 AA; 53145 MW; 6CC02052A86057BD CRC64;

Query Match
Best Local Similarity 93.6%; Score 422; DB 1; Length 507;
Matches 84; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAERAKAIAIGKGGEDTDPYKGVTFITDITKTGDDNGKYSTTINGEKVTLTVAD 60
DB 255 STAGTAERAKAIAIGKGGEDTDPYKGVTFITDITKTGDDNGKYSTTINGEKVTLTVAD 314

QY 61 IATGATDVAAATLQSSKNVYTSVVGQFTF 90
DB 315 IATGATDVAAATLQSSKNVYTSVVGQFTF 344

RESULT 8
FLIC_SALBE STANDARD; PRT; 504 AA.
AC 006970;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC.
OS Salmoneilla derby.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneilla.
OX NCBI_TaxID=28144;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-S241;
RX MEDLINE=94195780; PubMed=8146152;
RA Li J., Nelson K., McWhorter A.C., Whitlam T.S., Selander R.K.,
RT "Recombinational basis of serovar diversity in Salmoneilla enterica.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:2552-2556(1994).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 6960;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.U., Jovs T.M.;
RT "Molecular analyses of the Salmoneilla g. . . flagellar antigen
complex.";
RL J. Bacteriol. 175:5359-5365(1993).
RN (3)
RP FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SARMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U06225; AAA17866.1; -
CC EMBL; Z15066; CAA78773.1; -
CC InterPro; IPR001029; Flagellin_C.
CC InterPro; IPR001492; Flagellin_N.
CC Pfam; PF00700; Flagellin_C_1.
CC Pfam; PF00669; Flagellin_N_1.
DR PRINTS; PRO0207; FLAGELLIN.
KW Flagella.
FT INIT MET 0 0 BY SIMILARITY.
SEQUENCE 504 AA; 52818 MW; E428BB7FFBDA790 CRC64;

Query Match
Best Local Similarity 93.3%; Score 420; DB 1; Length 504;
Matches 84; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 STAGTAERAKAIAIGKGGEDTDPYKGVTFITDITKTGDDNGKYSTTINGEKVTLTVAD 60


```
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLIN, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL; Z15070; CAA78779.1; -

DR PIR; A53465; A53465.

DR PIR; S33192; S33192.

DR InterPro; IPR001029; Flagellin_C.

DR InterPro; IPR001492; Flagellin_N.

DR Pfam; PF00700; Flagellin_C; 1.

DR Pfam; PF00669; Flagellin_N; 1.

DR PRINTS; PR00207; FLAGELLIN.

KM Flagella.

FT INIT MET

SO SEQUENCE 507 AA; 53065 MW; 17A88B4BBA4A6125 CRC64;

Query Match

Best Local Similarity 80.7%; Score 364; DB 1; Length 507;

Matches 71; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 STAGTAAKATAGATGKGGEDTPTDKGTGDDGNGKVTINGEKYTLTVAD 60

DB 255 SAAGTDDAKAATATSKGKGVGDTPTDKGVSFTTIDKAGDGGVSTINGEKYTLTVAD 314

QY 61 IATGATDVNAATLQSSKNVYTSVNGQPTF 90

DB 315 IGASATDVNSAKIQSSKDVYTSVNGQPTF 344

RESULT 12

AL54 CANAL STANDARD; PRT; 469 AA.

AC 074660;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DS Agglutinin-like protein 4 precursor (Fragment).

GN AL54.

OS Candida albicans (Yeast).

CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetales; mitospotic Saccharomycetales; Candida.

OX NCBI_TaxID=5476;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1161;

RX MEDLINE=98440424; PubMed=9765564;

RA Hoyer L.L., Payne T.L., Hecht J.E.;

RT "Identification of Candida albicans ALS2 and ALS4 and localization of

RT a proteins to the fungal cell surface.";

RL J. Bacteriol. 180:5334-5343(1998).

CC -1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.

CC -1- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AF024584; AAC64239.1; -

KM Cell adhesion; Glycoprotein; Repeat; Signal.

FT SIGNAL 1 17

FT CHAIN 18 >469 AGGLUTININ-LIKE PROTEIN 4.

FT NON TER 469 469

SO SEQUENCE 469 AA; 49604 MW; 0BDCAB19889FBC1 CRC64;

Query Match

Best Local Similarity 29.4%; Score 82.5; DB 1; Length 469;

Matches 30; Conservative 17; Mismatches 38; Indels 17; Gaps 3;

QY 1 STAGTAAKATAGATGKGGEDTPTDKGTGDDGNGKVTINGEKYTLTVAD 51

DB 115 TTMTADTKAL-GVTTTPSPSVGSGSDVLANSCQCTAGINVTTFNDGDTSTVDF 173

QY 52 EKXVTLTVAD-----IATGATDVNAATLQSSKNVYTSVNG 86

DB 174 EKSTVASSDRLLRLRLPSLSQAVNLFPLPCANGYISGTMG 215

RESULT 13

FLB1 CAMJE STANDARD; PRT; 571 AA.

AC P56964; O9PMW1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DS Flagellin B.

GN FLAB OR CUI338C.

OS Campylobacter jejuni.

CC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;

CC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=197;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCYC 11168;

RX MEDLINE=20150912; PubMed=10688204;

RA Parkhill J., Wren B.W., Mungall K., Kertley J.M., Churcher C.,

RA Baham D., Chillingworth T., Davies R.M., Felwell T., Holtz S.,

RA Jagers K., Karlyshev A.V., Moule S., Pallen K.J., Penn C.W.,

RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrall B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

RT reveals hypervariable sequences.";

RL Nature 403:665-668(2000).

CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO

CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.

CC -1- SUBUNIT: HETEROPOLYMER OF FLAA AND FLAB.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AL139078; CAB73765.1; -

DR PIR; G81277; G81277.

DR InterPro; IPR001029; Flagellin_C.

DR InterPro; IPR001492; Flagellin_N.

DR Pfam; PF00700; Flagellin_C; 1.

DR Pfam; PF00669; Flagellin_N; 1.

DR PRINTS; PR00207; FLAGELLIN.

KM Flagella; Complete proteome.

FT INIT MET

SO SEQUENCE 571 AA; 59054 MW; ACE2D9B30CF05053 CRC64;

Query Match

Best Local Similarity 18.3%; Score 82.5; DB 1; Length 571;

Matches 25; Conservative 7; Mismatches 36; Indels 5; Gaps 2;

QY 5 TAAKATAGATGKGGEDTPTDKGTGDDGNGKVTINGEKYTLTVAD 64


```

Db      234 TVETRGIA-AVRAGTSDTFPAINGVTITGVAVEDGONGALVAINISVDTTGV-----EA 288
QY      65 ATDVNAATLOSSK 77
Db      289 SIDANGQLLTSR 301

RESULT 14
SLAP_CAMPE
ID      SLAP_CAMPE      STANDARD;      PRT;      933 AA.
AC      P55827;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      01-OCT-1994 (Rel. 30, Last annotation update)
DE      S-layer protein (surface array protein) (SAP).
GN      SARA.
OS      Campylobacter fetus.
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC      Campylobacteraceae; Campylobacter.
OX      NCBI_TaxId=196;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC      STRAIN=84-32 / 23D;
RX      MEDLINE=90354448; PubMed=2387868;
RA      Blaser M.J., Gotschlich E.C.;
RT      "Surface array protein of Campylobacter fetus. Cloning and gene
      structure";
RL      J. Biol. Chem. 265:14529-14535(1990).
RN      [2]
RP      ERRATUM.
RX      MEDLINE=91035477; PubMed=2229082;
RA      Blaser M.J., Gotschlich E.C.;
RL      J. Biol. Chem. 265:19372-19372(1990).
CC      -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
      OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
      CRITICAL FOR VIRULENCE.
CC      -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
      S-LAYER WITH HEXAGONAL SYMMETRY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J05577; AAA23032.1; -.
KW      Cell wall; S-layer.
SQ      SEQUENCE 933 AA; 96757 MW; F88C729B4BA5B1E9 CRC64;

Query Match
Best Local Similarity 18.3%; Score 82.5; DB 1; Length 933;
Matches 31; Conservative 9; Mismatches 40; Indels 21; Gaps 6;

QY      5 TAEAKAIA-----GAIKGKSGDTPDYKGYTF-TIDTKGDD-----GNGKVS-TTIN-- 50
Db      776 TPDASAITIGANATIKGSGGADSTIVKGNIVDLVAGSDPTTLTKKGAEXTDITTVNMF 835

QY      51 --GEKVTITVADIATGATDVNAATLOSSKRV--YTSVVG 86
Db      836 NAGDKT--DIADAKNGTFTFNKTIWMSDANLDDYITKAVAG 874

RESULT 15
FLAB_CAMOB
ID      FLAB_CAMOB      STANDARD;      PRT;      571 AA.
AC      P56963; Q9PMMO;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Flagellin A.

```

```

GN      FLAA OR CU1339C.
OS      Campylobacter jejuni.
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC      Campylobacteraceae; Campylobacter.
OX      NCBI_TaxId=197;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NCTC 11168;
RX      MEDLINE=20150912; PubMed=10689204;
RA      Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA      Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA      Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA      Quail M.A., Rastreadm M.A., Rutherford K.M., van Vliet A.H.M.,
RA      Whitehead S., Barrrell B.G.;
RT      "The genome sequence of the food-borne pathogen Campylobacter jejuni
      reveals hypervariable sequences.";
RL      Nature 403:665-668(2000).
CC      -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
      FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC      -!- SUBUNIT: HETEROPOLYMER OF FLAA AND FLAB.
CC      -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AL139078; CAB73766.1; -.
DR      PIR; H61277; H61277.
DR      InterPro; IPR001029; Flagellin_C.
DR      InterPro; IPR001492; Flagellin_N.
DR      Pfam; PF00700; Flagellin_C_1.
DR      Pfam; PF00669; Flagellin_N; 1.
DR      PRINTS; PR00207; Flagellin.
DR      ProDom; PD000316; Flagellin_C; 2.
KW      Flagella; Complete proteome.
RN      [1]
RT      INTER-MER
SQ      SEQUENCE 571 AA; 58907 MW; 9C63B2A10C1AE863 CRC64;

Query Match
Best Local Similarity 17.4%; Score 78.5; DB 1; Length 571;
Matches 26; Conservative 8; Mismatches 33; Indels 7; Gaps 4;

QY      5 TAEAKAIAIKGKSGDTPDYKGYTF-TIDTKGDDGNGKVSITINGEKVTLTVADIAT 63
Db      234 TVETRGIA-AVRAGTSDTFPAINGVKIGKVDYKDG-DANGALVAINISVDTTGV-----E 287

QY      64 GATDVNAATLOSSK 77
Db      288 ASIDANGQLLTSR 301

```

Search completed: January 12, 2004, 11:17:47
 Job time : 20 secs

FILE REFERENCE: 1038-861 MIS:jb
CURRENT APPLICATION NUMBER: US/09/206,942
CURRENT FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 09/167,568
EARLIER FILING DATE: 1998-10-07
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 37
LENGTH: 1222
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-206-942-37

Query Match 20.3%; Score 91.5; DB 4; Length 1222;
Best Local Similarity 24.3%; Pred. No. 0.052; 37; Indels 57; Gaps 5;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;

QY 1 STAGTAFAKAIAGAIKGE-----GDPDYKGV----- 30
DB 717 ATGSEVETAKTGIDKIGESNGVNITASGDTLVNSITGQNTVVAASGAVTTTKGS 776
QY 31 -----FTIDTKGD-----DNGKYSTT-----INGEYTLTV---ADIAT 63
DB 777 TINATTGANITTTKGEINGEVKSAGNVNITASGNTLVNSITGQNTVTVANSAGAITTT 836
QY 64 GATDVNAATLQSSKNVYTSVNGQ 87
DB 837 EGSTINATTGDANITTTGTGNINCK 860

RESULT 3
US-09-206-942-34
Sequence 34, Application US/09206942
Patent No. 6432669

GENERAL INFORMATION:
APPLICANT: Loomore, Sheena M.
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
FILE REFERENCE: 1038-861 MIS:jb
CURRENT APPLICATION NUMBER: US/09/206,942
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 09/167,568
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 34
LENGTH: 1228
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-206-942-34

Query Match 20.3%; Score 91.5; DB 4; Length 1228;
Best Local Similarity 24.3%; Pred. No. 0.053;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;

QY 1 STAGTAFAKAIAGAIKGE-----GDPDYKGV----- 30
DB 723 ATGSEVETAKTGIDKIGESNGVNITASGDTLVNSITGQNTVVAASGAVTTTKGS 782
QY 31 -----FTIDTKGD-----DNGKYSTT-----INGEYTLTV---ADIAT 63
DB 783 TINATTGANITTTKGEINGEVKSAGNVNITASGNTLVNSITGQNTVTVANSAGAITTT 842
QY 64 GATDVNAATLQSSKNVYTSVNGQ 87
DB 843 EGSTINATTGDANITTTGTGNINCK 866

RESULT 4
US-09-206-942-28
Sequence 28, Application US/09206942

Patent No. 6432669
GENERAL INFORMATION:
APPLICANT: Loomore, Sheena M.
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
FILE REFERENCE: 1038-861 MIS:jb
CURRENT APPLICATION NUMBER: US/09/206,942
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 09/167,568
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 28
LENGTH: 1220
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-206-942-28

Query Match 20.1%; Score 90.5; DB 4; Length 1220;
Best Local Similarity 24.3%; Pred. No. 0.069;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;

QY 1 STAGTAFAKAIAGAIKGE-----GDPDYKGV----- 30
DB 715 ATGSEVETAKTGIDKIGESNGVNITASGDTLVNSITGQNTVVAASGAVTTTKGS 774
QY 31 -----FTIDTKGD-----DNGKYSTT-----INGEYTLTV---ADIAT 63
DB 775 TINATTGANITTTKGEINGEVKSAGNVNITASGNTLVNSITGQNTVTVANSAGAITTT 834
QY 64 GATDVNAATLQSSKNVYTSVNGQ 87
DB 835 EGSTINATTGDANITTTGTGNINCK 858

RESULT 5
US-09-206-942-26
Sequence 26, Application US/09206942
Patent No. 6432669

GENERAL INFORMATION:
APPLICANT: Loomore, Sheena M.
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
FILE REFERENCE: 1038-861 MIS:jb
CURRENT APPLICATION NUMBER: US/09/206,942
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 09/167,568
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 26
LENGTH: 1226
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-206-942-26

Query Match 20.1%; Score 90.5; DB 4; Length 1226;
Best Local Similarity 24.3%; Pred. No. 0.069;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;

QY 1 STAGTAFAKAIAGAIKGE-----GDPDYKGV----- 30
DB 721 ATGSEVETAKTGIDKIGESNGVNITASGDTLVNSITGQNTVVAASGAVTTTKGS 780
QY 31 -----FTIDTKGD-----DNGKYSTT-----INGEYTLTV---ADIAT 63
DB 781 TINATTGANITTTKGEINGEVKSAGNVNITASGNTLVNSITGQNTVTVANSAGAITTT 840
QY 64 GATDVNAATLQSSKNVYTSVNGQ 87

Db 841 EGSTINATTTGDANITTTGTGNGK 864

RESULT 6

US-08-728-470-9
; Sequence 9, Application US/08728470
; Patent No. 5928651
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,470
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstreaser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-728-470-9

Query Match 19.2%; Score 86.5; DB 2; Length 1338;
Best Local Similarity 25.4%; Pred. No. 0.23;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAAKAIAAGIKGKE-----GDTFDYKGYT----- 30
Db 936 ATSGTVNISRTGDKIGISTSGNVNITASGNLTKVSNITGDVTVTADAGALTTTAA 995
QY 31 -----FTIDFTKDDGNGKYSTTINGEKVLTVDIATGATDVAAATLQSSKNVYTS 82
Db 996 TISATTGNANITTKTSDI-NKGVSSSGSVTLVATGATLAVGNISGTVTTITADSGKLT 1054
QY 83 VV 84
Db 1055 TV 1056

RESULT 7
US-08-719-641-9

; Sequence 9, Application US/08719641
; Patent No. 6218141
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 6218141-Typeable Haemophilus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/719,641
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstreaser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-625
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-719-641-9

Query Match 19.2%; Score 86.5; DB 3; Length 1338;
Best Local Similarity 25.4%; Pred. No. 0.23;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAAKAIAAGIKGKE-----GDTFDYKGYT----- 30
Db 936 ATSGTVNISRTGDKIGISTSGNVNITASGNLTKVSNITGDVTVTADAGALTTTAA 995
QY 31 -----FTIDFTKDDGNGKYSTTINGEKVLTVDIATGATDVAAATLQSSKNVYTS 82
Db 996 TISATTGNANITTKTSDI-NKGVSSSGSVTLVATGATLAVGNISGTVTTITADSGKLT 1054
QY 83 VV 84
Db 1055 TV 1056

RESULT 8
US-08-728-470-10
; Sequence 10, Application US/08728470
; Patent No. 5928651
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,470
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Berktessser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-633
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-728-470-10

Query Match 19.2%; Score 86.5; DB 2; Length 1529;
Best Local Similarity 25.4%; Pred. No. 0.28;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAEAKAIAAGKIGKE-----GDFDYKGV----- 30
DB 1127 ATSGTWNISTKGTGIGKIGESTSGNWNITASGNLTAKSNITGQDVTVADAGALTTAGS 1166
QY 31 -----FTIDTKGDDGNGKVSSTINGEKVTLTVADIATGATDVNATLQSSKNVTS 82
DB 1187 TISATGNNATITTKGDI-NKGVSSSGSVTLVATGATLAVGNISGNTVTTADSGKLT 1245

QY 83 VV 84
DB 1246 TV 1247

RESULT 9
US-08-719-641-10
Sequence 10, Application US/08719641
Patent No. 6218141
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 6218141-Typeable Haemophilus
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
CITY: Arlington

STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,641
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Berktessser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-625
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-719-641-10

Query Match 19.2%; Score 86.5; DB 3; Length 1529;
Best Local Similarity 25.4%; Pred. No. 0.28;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAEAKAIAAGKIGKE-----GDFDYKGV----- 30
DB 1127 ATSGTWNISTKGTGIGKIGESTSGNWNITASGNLTAKSNITGQDVTVADAGALTTAGS 1166
QY 31 -----FTIDTKGDDGNGKVSSTINGEKVTLTVADIATGATDVNATLQSSKNVTS 82
DB 1187 TISATGNNATITTKGDI-NKGVSSSGSVTLVATGATLAVGNISGNTVTTADSGKLT 1245

QY 83 VV 84
DB 1246 TV 1247

RESULT 10
US-08-617-697-9
Sequence 9, Application US/08617697
Patent No. 5877336
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 5877336-Typeable Haemophilus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
STREET: Bldg. 1
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,697
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 05-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Beikstreser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-557
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1599 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-697-9

Query Match 19.2%; Score 86.5; DB 2; Length 1599;
Best Local Similarity 25.4%; Pred. No. 0.3;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAARAKAIAGIKGKE-----GDTFDYKGVY----- 30
DB 1196 ATSGTVNISTKTGDIKIGIESTGNNVITASGNTLAKVSNITGQDVVTADAGALTITAGS 1255

QY 31 -----FTIDKTDGDDGKXVSTTINGEKVTLTVADITGATDVNAATLQSSKNVYTS 82
DB 1256 TISATGNNANITTKTGDI-NGKVESSSGSVTLVATGATLAVGNISGNTVITADSGKLTIS 1314

QY 83 VV 84
DB 1315 TV 1316

RESULT 11
US-08-617-697-10
Sequence 10; Application US/08617697
Patent No. 5977336
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Shoemaker and Matzare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,697
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,832
FILING DATE: 05-OCT-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Beikstreser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-557
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-697-10

Query Match 19.2%; Score 86.5; DB 2; Length 1600;
Best Local Similarity 25.4%; Pred. No. 0.3;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAARAKAIAGIKGKE-----GDTFDYKGVY----- 30
DB 1197 ATSGTVNISTKTGDIKIGIESTGNNVITASGNTLAKVSNITGQDVVTADAGALTITAGS 1256

QY 31 -----FTIDKTDGDDGKXVSTTINGEKVTLTVADITGATDVNAATLQSSKNVYTS 82
DB 1257 TISATGNNANITTKTGDI-NGKVESSSGSVTLVATGATLAVGNISGNTVITADSGKLTIS 1315

QY 83 VV 84
DB 1316 TV 1317

RESULT 12
US-08-409-995-4
Sequence 4; Application US/08409995
Patent No. 5646259
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen I.
TITLE OF INVENTION: St. Gene III, Joseph W.
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Flehr, Hobach, Test, Albrighton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/409,995
FILING DATE: 24-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-61053/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: unknown
US-08-409-995-4

Query Match 19.0%; Score 85.5; DB 1; Length 1912;
Best Local Similarity 39.1%; Pred. No. 0.5;
Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAAKALAGATGCGEGDTPDYKVTFTIDTKTGDDGNGKSVTTINGEKVTLTVADIA 62
DB 1435 SATVSDKSLGT-NGNKVNITSDTKGLNFAKSKTGDDAN-----IHNGIASLTITDITL 1489
QY 63 TGAT 66
DB 1490 SGAT 1493

RESULT 13
US-08-685-467-4
Sequence 4, Application US/08685467
Patent No. 6060059

GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
ATTORNEY/AGENT INFORMATION: Barenkamp, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobdach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,467
FILING DATE: 22-JUL-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-685-467-4

Query Match 19.0%; Score 85.5; DB 3; Length 1912;
Best Local Similarity 39.1%; Pred. No. 0.5;
Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAAKALAGATGCGEGDTPDYKVTFTIDTKTGDDGNGKSVTTINGEKVTLTVADIA 62
DB 1435 SATVSDKSLGT-NGNKVNITSDTKGLNFAKSKTGDDAN-----IHNGIASLTITDITL 1489
QY 63 TGAT 66
DB 1490 SGAT 1493

RESULT 14
US-09-377-155-33

Sequence 33, Application US/09377155
Patent No. 6197312
GENERAL INFORMATION:
APPLICANT: PEAK, Ian Richard Anselm
ATTORNEY/AGENT INFORMATION: JENNINGS, Michael Paul
TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
FILE REFERENCE: 065064/0128
CURRENT APPLICATION NUMBER: US/09/377,155
CURRENT FILING DATE: 1999-08-19
PRIOR APPLICATION NUMBER: PCT/AU98/01031
PRIOR FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: GB 9726398.2
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 33
LENGTH: 2353
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-377-155-33

Query Match 19.0%; Score 85.5; DB 3; Length 2353;
Best Local Similarity 39.1%; Pred. No. 0.67; Mismatches 28; Indels 5; Gaps 2;
Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAAKALAGATGCGEGDTPDYKVTFTIDTKTGDDGNGKSVTTINGEKVTLTVADIA 62
DB 1436 SATVSDKSLGT-NGNKVNITSDTKGLNFAKSKTGDDAN-----IHNGIASLTITDITL 1490
QY 63 TGAT 66
DB 1491 SGAT 1494

RESULT 15
US-08-913-942-4
Sequence 4, Application US/08913942
Patent No. 6200578
GENERAL INFORMATION:
APPLICANT: St. Gene, Joseph
ATTORNEY/AGENT INFORMATION: Barenkamp, Stephen J.
TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobdach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,942
FILING DATE: 29-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,995
FILING DATE: 24-MAR-1995
PRIOR APPLICATION NUMBER: PCT/US96/4031
FILING DATE: 22-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vance, Dolly A.
REGISTRATION NUMBER: 39,054
REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DAY

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277239
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2353 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-913-942-4

Query Match 19.0%; Score 85.5; DB 3; Length 2353;
Best Local Similarity 39.1%; Pred. No. 0.67;
Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAAKAIAGAIKGGKSDTFDYKGVFTIDTKTGDDGNGKVSITTINGEKVTLTVADIA 62
Db 1436 SATVSDKLSLST-NGNKVNITSDTKGINFAKDSKTGDDN---IHNGIASLTITDTLIN 1490
QY 63 TGAT 66
Db 1491 SGAT 1494

Search completed: January 12, 2004, 11:19:58
Job time: 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 11:18:26 ; Search time 44 Seconds
(without alignments) 412.179 Million cell updates/sec

Title: US-10-019-676-2
Sequence: 1 STACTAEAKAIAIKGKKE.....ATLSSKRVYTSVNGQTF 90

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91.5	20.3	915	12	US-10-193-764-35 Sequence 35, App1
2	91.5	20.3	1222	12	US-10-193-764-37 Sequence 37, App1
3	91.5	20.3	1228	12	US-10-193-764-34 Sequence 34, App1
4	90.5	20.1	1220	12	US-10-193-764-28 Sequence 28, App1
5	90.5	20.1	1226	12	US-10-193-764-26 Sequence 26, App1
6	86.5	19.2	1599	14	US-10-092-880-9 Sequence 9, App1
7	86.5	19.2	1600	14	US-10-092-880-10 Sequence 10, App1
8	85.5	19.0	2353	10	US-09-797-862-33 Sequence 33, App1
9	82.5	18.3	469	15	US-10-245-802-14 Sequence 14, App1
10	82.5	18.3	1073	12	US-10-193-764-45 Sequence 45, App1
11	82.5	18.3	1079	12	US-10-193-764-43 Sequence 43, App1
12	79	17.5	1005	12	US-10-193-764-41 Sequence 41, App1
13	79	17.5	1011	12	US-10-193-764-39 Sequence 39, App1
14	78.5	17.4	572	15	US-10-125-692-11 Sequence 11, App1
15	77.5	17.2	394	15	US-10-125-692-14 Sequence 14, App1

16	76.5	17.0	977	12	US-10-193-764-49 Sequence 49, App1
17	76.5	17.0	983	12	US-10-193-764-47 Sequence 47, App1
18	76.5	17.0	1004	12	US-10-193-764-53 Sequence 53, App1
19	76.5	17.0	1010	12	US-10-193-764-51 Sequence 51, App1
20	76.5	16.9	1180	12	US-10-193-764-61 Sequence 61, App1
21	76.5	16.9	1188	12	US-10-193-764-59 Sequence 59, App1
22	76.5	16.9	1477	14	US-10-092-880-4 Sequence 4, App1
23	75.5	16.7	626	9	US-09-765-272-106 Sequence 106, App1
24	75.5	16.7	1741	12	US-09-971-536-68 Sequence 68, App1
25	75.5	16.6	992	12	US-10-193-764-57 Sequence 57, App1
26	74.5	16.5	998	12	US-10-193-764-55 Sequence 55, App1
27	74.5	16.5	1098	10	US-09-797-862-32 Sequence 32, App1
28	74.5	16.4	185	10	US-09-764-847-516 Sequence 516, App1
29	74.5	16.4	185	10	US-10-092-154-516 Sequence 516, App1
30	74.5	16.4	1325	9	US-09-741-663-304 Sequence 304, App1
31	72.5	16.0	1036	12	US-10-193-764-69 Sequence 69, App1
32	72.5	16.0	1477	12	US-10-193-764-67 Sequence 67, App1
33	71.5	15.7	172	12	US-10-028-248A-92 Sequence 92, App1
34	71.5	15.7	1248	10	US-09-738-626-4814 Sequence 4814, App1
35	70.5	15.6	379	12	US-10-369-453-10459 Sequence 10459, App1
36	70.5	15.6	599	10	US-09-797-862-15 Sequence 15, App1
37	70.5	15.6	599	10	US-09-797-862-15 Sequence 15, App1
38	70.5	15.6	3871	15	US-10-184-644-347 Sequence 347, App1
39	70.5	15.6	3871	15	US-10-184-644-347 Sequence 347, App1
40	70.5	15.5	569	15	US-10-156-761-13756 Sequence 13756, App1
41	70.5	15.5	621	10	US-09-738-626-3937 Sequence 3937, App1
42	69.5	15.4	524	11	US-09-847-208-110 Sequence 110, App1
43	69.5	15.4	524	11	US-09-791-279-105 Sequence 105, App1
44	69.5	15.3	528	12	US-10-021-660-73 Sequence 73, App1
45	69.5	15.3	595	10	US-09-738-626-5469 Sequence 5469, App1

ALIGNMENTS

RESULT 1
US-10-193-764-35
Sequence 35, Application US/10193764
Publication No. US20030133943A1
GENERAL INFORMATION:
APPLICANT: Loesmore, Sheena M.
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
FILE REFERENCE: 1038-1239MIS
CURRENT APPLICATION NUMBER: US/10/193,764
CURRENT FILING DATE: 2002-07-12
PRIOR APPLICATION NUMBER: 09/167,568
NUMBER OF SEQ ID NOS: 91
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 915
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-10-193-764-35

Query Match 20.3%, Score 91.5; DB 12; Length 915;
Best Local Similarity 24.3%, Pred. No. 0.16;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;
QY 1 STACTAEAKAIAIKGKKE-----GDFDYKGV----- 30
DB 410 ATTGSVVTAKTGIGKIGIESGNVITIGQVTVAAAGAVTTKGS 469
QY 31 -----FTIDTGTGPD-----DNGKVSST----- 63
DB 470 TTAATGNANITTTGGEINEERVSASGNVITIASGNTLANSITIGQVTVANSAGATTT 529
QY 64 GATDVNATLSSKRVYTSVNGQ 87
DB 530 EGTINATGTGANITTTGTGNIKG 553

RESULT 2

US-10-193-764-37
; Sequence 37, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Klein, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-37

Query Match

Best Local Similarity 20.3%; Score 91.5; DB 12; Length 1222;
Best Local Similarity 24.3%; Pred. No. 0.24; Mismatches 37; Indels 57; Gaps 5;
Matches 35; Conservative 15;

Query

1 STAGTAEAKAIAGAIKGGKE-----GDTFDYKGV----- 30
717 ATGSGVEVTKAKTDIKGIEBSNGVNTASGDTLVNSNITGQNVTVAAASGAVTTTGS 776

Db

31 -----FTIDRTGD-----DGNKXSTT-----INGEKVTLTV---ADIAT 63
777 TINATTGNANITTKTGIEBSNGVNTASGDTLVNSNITGQNVTVAAASGAVTTT 836

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
837 EGSTINATTGDANITTTQTGNINCK 860

Query 31 -----FTIDRTGD-----DGNKXSTT-----INGEKVTLTV---ADIAT 63
Db 783 TINATTGNANITTKTGIEBSNGVNTASGDTLVNSNITGQNVTVAAASGAVTTTGS 842
Query 64 GATDVNAATLQSSKNVYTSVNGQ 87
Db 843 EGSTINATTGDANITTTQTGNINCK 866

RESULT 4

US-10-193-764-28
; Sequence 28, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Klein, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-28

Query Match

Best Local Similarity 20.1%; Score 90.5; DB 12; Length 1220;
Best Local Similarity 24.3%; Pred. No. 0.31; Mismatches 37; Indels 57; Gaps 5;
Matches 35; Conservative 15;

Query

1 STAGTAEAKAIAGAIKGGKE-----GDTFDYKGV----- 30
715 ATGSGVEVTKAKTDIKGIEBSNGVNTASGDTLVNSNITGQNVTVAAASGAVTTTGS 774

Db

31 -----FTIDRTGD-----DGNKXSTT-----INGEKVTLTV---ADIAT 63
775 TINATTGNANITTKTGIEBSNGVNTASGDTLVNSNITGQNVTVAAASGAVTTT 834

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Db

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Query

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

64 GATDVNAATLQSSKNVYTSVNGQ 87
835 EGSTINATTGDANITTTQTGNINCK 858

Best Local Similarity 24.3%; Pred. No. 0.31;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;
QY 1 STAGTAEAKAIAAGKGGKE-----GDFPYKGV----- 30
DB 721 ATIGSEVETAKTGDIKGGIESNSGNVNITASGTLVNSITGQNTVVAASGAVTTTKGS 780
QY 31 -----FTIDTKGDDGNGKVSSTINGEKVTLVADIATGATDVNAATLQSSKNVYTS 82
DB 781 TINATGNANITTKTEINGEYKASAGNVNITASGTLVNSITGQNTVITANSGLITTT 840
QY 64 GADVNAATLQSSKNVYTSVWNCQ 87
DB 841 EGSTINATGDANITQTGTINCK 864

RESULT 6
US-10-092-880-9
Sequence 9, Application US/10092880
Publication No. US20020164354A1
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J.
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS OF NON-TYPEABLE
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/092,880
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: 09/155,614
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 08/617,697
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: PCT/US97/04707
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 9
LENGTH: 1599
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-10-092-880-9

Query Match 19.2%; Score 86.5; DB 14; Length 1599;
Best Local Similarity 25.4%; Pred. No. 1.2;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;
QY 1 STAGTAEAKAIAAGKGGKE-----GDFPYKGV----- 30
DB 1196 ATSGVNI STKTGDIKGGIESGTVNITASGTLKVSNTIGQDVTVTADAGALTTTAGS 1255
QY 31 -----FTIDTKGDDGNGKVSSTINGEKVTLVADIATGATDVNAATLQSSKNVYTS 82
DB 1256 TISATGNANITTKGDI-NKGVSSSGSVTLVATGATLAVGNISGNTVITADSGKLTS 1314
QY 83 VV 84
DB 1315 TV 1316

RESULT 7
US-10-092-880-10
Sequence 10, Application US/10092880
Publication No. US20020164354A1
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J.
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS OF NON-TYPEABLE
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/092,880
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: 09/155,614
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 08/617,697
PRIOR FILING DATE: 1996-04-01

PRIOR APPLICATION NUMBER: PCT/US97/04707
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 10
LENGTH: 1600
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-10-092-880-10

Query Match 19.2%; Score 86.5; DB 14; Length 1600;
Best Local Similarity 25.4%; Pred. No. 1.2;
Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAEAKAIAAGKGGKE-----GDFPYKGV----- 30
DB 1197 ATSGVNI STKTGDIKGGIESGTVNITASGTLKVSNTIGQDVTVTADAGALTTTAGS 1256
QY 31 -----FTIDTKGDDGNGKVSSTINGEKVTLVADIATGATDVNAATLQSSKNVYTS 82
DB 1257 TISATGNANITTKGDI-NKGVSSSGSVTLVATGATLAVGNISGNTVITADSGKLTS 1315
QY 83 VV 84
DB 1316 TV 1317

RESULT 8
US-09-797-862-33
Sequence 33, Application US/09797862
Patent No. US20020102276A1
GENERAL INFORMATION:
APPLICANT: PEAK, IAN RICHARD ANSELM
APPLICANT: JENNINGS, MICHAEL PAUL
APPLICANT: MOXON, E. RICHARD
TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
FILE REFERENCE: 065064/0134
CURRENT APPLICATION NUMBER: US/09/797,862
CURRENT FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: PCT/AU98/01031
PRIOR FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: GB 9726398.2
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 33
LENGTH: 2353
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-797-862-33

Query Match 19.0%; Score 85.5; DB 10; Length 2353;
Best Local Similarity 39.1%; Pred. No. 2.6;
Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAEAKAIAAGKGGKEGDFPYKGVFTIDTKGDDGNGKVSSTINGEKVTLVADIA 62
DB 1436 SATVSDKLSLGT-NGKNVNTITSDYKGLNFAKDSKTGDAN-----IHNGIASITLDTILN 1490
QY 63 TGAT 66
DB 1491 SGAT 1494

RESULT 9
US-10-245-802-14
Sequence 14, Application US/10245802
Publication No. US20030124134A1
GENERAL INFORMATION:
APPLICANT: Edwards, John E.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND METHODS TO VACCINATE AGAINST
FILE REFERENCE: 013361.4003

;; CURRENT APPLICATION NUMBER: US/10/245,802
;; CURRENT FILING DATE: 2002-09-13
;; PRIOR APPLICATION NUMBER: US 09/715,876
;; PRIOR FILING DATE: 2000-11-18
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patent In version 3.2
;; SEQ ID NO 14
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Candida albicans
US-10-245-802-14

Query Match 18.3%; Score 82.5; DB 15; Length 469;
Best Local Similarity 29.4%; Pred. No. 0.69;
Matches 30; Conservative 17; Mismatches 38; Indels 17; Gaps 3;

QY 1 STAGTAARAKAIAGAIK-----GKRGDPDYKGVFTTIDTKGDDGNGKV-----STTINGKVT 51
DB 115 TTMTADTKAI-GVTLTLPFSVSGSDVDLANSQCFAGINTVTFNDGDTISITVDF 173

QY 52 EKVTLTVAD-----IATGATDVNAATLQSSKNVTSVNG 86
DB 174 EKSTVASSDRLLSRILPSLSQAVNLPFLPQECANGTSGTW 215

RESULT 10
US-10-193-764-45
; Sequence 45, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1238MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-45

Query Match 18.3%; Score 82.5; DB 12; Length 1073;
Best Local Similarity 26.5%; Pred. No. 2;
Matches 27; Conservative 18; Mismatches 36; Indels 21; Gaps 4;

QY 1 STAGTAARAKAIAGAIKRGKGDTPDYKGVFTTIDTKGDDGNGKV-----STTINGKVT 55
DB 714 ATTGNVLTAVTSDIQGIRKNSGD-----VNITTSIGSI-NGKIESKSGSVTLTAIEKT 767

QY 56 LTVADIATGATDVNA-----ATLQSSKNVTSVNGQ 87
DB 768 LTVGNVSGNTVTVTANRGALTTLAGSTINGTNGVTSSQSGE 809

RESULT 11
US-10-193-764-43
; Sequence 43, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764

;; CURRENT FILING DATE: 2002-07-12
;; PRIOR APPLICATION NUMBER: 09/167,568
;; PRIOR FILING DATE: 1998-10-07
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 43
;; LENGTH: 1079
;; TYPE: PRT
;; ORGANISM: Haemophilus influenzae
US-10-193-764-43

Query Match 18.3%; Score 82.5; DB 12; Length 1079;
Best Local Similarity 26.5%; Pred. No. 2;
Matches 27; Conservative 18; Mismatches 36; Indels 21; Gaps 4;

QY 1 STAGTAARAKAIAGAIKRGKGDTPDYKGVFTTIDTKGDDGNGKV-----STTINGKVT 55
DB 720 ATTGNVLTAVTSDIQGIRKNSGD-----VNITTSIGSI-NGKIESKSGSVTLTAIEKT 773

QY 56 LTVADIATGATDVNA-----ATLQSSKNVTSVNGQ 87
DB 774 LTVGNVSGNTVTVTANRGALTTLAGSTINGTNGVTSSQSGE 815

RESULT 12
US-10-193-764-41
; Sequence 41, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-41

Query Match 17.5%; Score 79; DB 12; Length 1005;
Best Local Similarity 23.0%; Pred. No. 4.6;
Matches 29; Conservative 15; Mismatches 38; Indels 44; Gaps 3;

QY 1 STAGTAARAKAIAGAIKRGKGDTPDYKGVFTTIDTK-----STTINGKVT 36
DB 704 ATTGSVEVYAKTGDISGTSIGKTVATATDSLTVGAKINATBGTATLTLSSGKLTTE 763

QY 37 -----TGDDNGKVTSTINGEKVTLVA-----DIATGATDVNAATLQSSK 77
DB 764 ANSAISGANGVTAASSGSGISGTSIGKTVSVTAASSSLVGVGDAKINATE-GAATLTATK 822

QY 78 NVYTSV 83
DB 823 GTLTTV 828

RESULT 13
US-10-193-764-39
; Sequence 39, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH

;; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
;; FILE REFERENCE: 1038-1239MIS
;; CURRENT APPLICATION NUMBER: US/10/193,764
;; CURRENT FILING DATE: 2002-07-12
;; PRIOR APPLICATION NUMBER: 09/167,568
;; PRIOR FILING DATE: 1998-10-07
;; NUMBER OF SEQ ID NOS: 91
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 39
;; LENGTH: 1011
;; TYPE: PRT
;; ORGANISM: Haemophilus influenzae
US-10-193-764-39

Query Match 17.5%; Score 79; DB 12; Length 1011;
Best Local Similarity 23.0%; Pred. No. 4.6;
Matches 29; Conservative 15; Mismatches 38; Indels 44; Gaps 3;

QY 1 STAGTAAKAIAGAIKGGKSGDTPYKGVTFITDK----- 36
DB 710 ATTSVEVYAKTGDISTGKTSVATTDLSLVKGGKINATEGATLASSGKLTTE 769
QY 37 -----TGDDGNGKVSSTINGEKVTLTVA-----DIATGATVNAATLQSSK 77
DB 770 ANSAISGANGVYASSQSGDISGTSKTVSVTASSGSLTVGGDAKINATE-GAATLTATK 828
QY 78 NVYTSV 83
DB 829 GTLTIV 834

RESULT 14
US-10-125-692-11
; Sequence 11, Application US/10125692
; Publication No. US20030044429A1
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan
; APPLICANT: Hayashi, Fumitaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozinsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 ligands and Methods
; TITLE OF INVENTION: of Use
; FILE REFERENCE: P-IS 5155
; CURRENT APPLICATION NUMBER: US/10/125,692
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/285,477
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 572
; TYPE: PRT
; ORGANISM: C. jejuni
US-10-125-692-11

Query Match 17.4%; Score 78.5; DB 15; Length 572;
Best Local Similarity 35.1%; Pred. No. 2.5;
Matches 26; Conservative 8; Mismatches 33; Indels 7; Gaps 4;

QY 5 TAAKAIAGAIKGGKSGDTPYKGVTF-TIDTKGDDGNGKVSSTINGEKVTLTVADIAT 63
DB 235 TVETTRGIA-AVRAGATSDTPAINGVKGKVDYKDG-DANGALVAALNSVKRTGV---E 288
QY 64 GATDVNAATLQSSK 77
DB 289 ASIDANGQLLTISR 302

RESULT 15
US-10-125-692-14
; Sequence 14, Application US/10125692
; Publication No. US20030044429A1

;; GENERAL INFORMATION:
;; APPLICANT: Aderem, Alan
;; APPLICANT: Hayashi, Fumitaka
;; APPLICANT: Smith, Kelly D.
;; APPLICANT: Underhill, David M.
;; APPLICANT: Ozinsky, Adrian
;; TITLE OF INVENTION: Toll-Like Receptor 5 ligands and Methods
;; TITLE OF INVENTION: of Use
;; FILE REFERENCE: P-IS 5155
;; CURRENT APPLICATION NUMBER: US/10/125,692
;; CURRENT FILING DATE: 2002-04-17
;; PRIOR APPLICATION NUMBER: US 60/285,477
;; PRIOR FILING DATE: 2001-04-20
;; NUMBER OF SEQ ID NOS: 43
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 394
;; TYPE: PRT
;; ORGANISM: P. aeruginosa
US-10-125-692-14

Query Match 17.2%; Score 77.5; DB 15; Length 394;
Best Local Similarity 32.5%; Pred. No. 2;
Matches 25; Conservative 12; Mismatches 19; Indels 21; Gaps 5;

QY 2 TAGTAAKAIAGAIKGG-----KEGTPYKGVTFITDKTGDDGNGKVSSTINGEKV 54
DB 214 TAEQAAK-IAAANDANVIGAFSDGDTISYV-----SKAGKDSGGAITSVSG--- 262
QY 55 TLTVADIATGATDVNAA 71
DB 263 -VVIAD--TGSTGVCTA 276

Search completed: January 12, 2004, 11:25:14
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 11:16:17 (Search time 47 seconds
(without alignments)
303.945 Million cell updates/sec

Title: US-10-019-676-2
Perfect score: 451

Sequence: 1 STAGTAEMAKATGAIKGGKE.....ATLSSKNVYTSVNGQFTF 90

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneeq/geneeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451	100.0	90	22	AA066796
2	446	98.9	493	22	AA031432
3	427	94.7	90	22	AA031433
4	349	77.4	69	22	AA031434
5	197	43.7	40	22	AA031435
6	133	29.5	27	22	AA031436
7	91.5	20.3	1222	21	AA01830
8	91.5	20.3	1228	21	AA01838
9	90	20.0	1221	21	AA01825

10	90	20.0	1227	21	AA01824
11	86.5	19.2	1338	14	AA041731
12	86.5	19.2	1529	14	AA041732
13	86.5	19.2	1598	18	AA030291
14	86.5	19.2	1601	18	AA030292
15	85.5	19.0	2353	17	AA093393
16	85.5	19.0	2411	21	AA023860
17	82.5	18.3	1073	21	AA023867
18	82.5	18.3	1079	21	AA023836
19	82	18.2	1874	19	AA064518
20	81.5	18.1	241	22	AA076197
21	81.5	18.1	241	22	AA050019
22	79	17.5	1005	21	AA01833
23	79	17.5	1011	21	AA01832
24	79	17.5	1386	24	AA082573
25	79	17.5	1449	24	AA082570
26	76.5	17.0	957	21	AA01839
27	76.5	17.0	963	21	AA01838
28	76.5	17.0	1004	21	AA01841
29	76.5	17.0	1010	21	AA01840
30	76	16.9	1094	21	AA023858
31	76	16.9	1180	21	AA01845
32	76	16.9	1188	21	AA01844
33	76	16.9	1477	14	AA041724
34	75.5	16.7	613	22	AA062232
35	75.5	16.7	626	19	AA055115
36	75.5	16.7	626	23	AA054609
37	75.5	16.7	857	24	AA020861
38	75.5	16.7	1741	23	AA020109
39	75	16.6	992	21	AA01843
40	75	16.6	998	21	AA01842
41	74.5	16.5	258	23	AA05962
42	74.5	16.5	258	23	AA05962
43	74.5	16.5	287	23	AA05982
44	74.5	16.5	1098	17	AA093392
45	74	16.4	185	22	AA06451

ALIGNMENTS

RESULT 1	AA066796	standard; Protein; 90 AA.
ID	AA066796	
AC	AA066796	
DT	12-APR-2001	(first entry)
XX	Part of <i>Salmonella enteritidis</i> flagellin protein.	
DE	Immunodiagnostic; piezoelectric crystal; frequency infection.	
XX		
KW	<i>Salmonella enteritidis</i> .	
OS	<i>Salmonella enteritidis</i> .	
XX		
PN	WC000102858-AL.	
XX		
PD	11-JAN-2001.	
XX		
FE	04-OCT-1999;	99WO-SG00098.
XX		
PR	05-JUL-1999;	99SG-0003147.
XX		
PA	(MOLE-) INST MOLECULAR AGROBIOLOGY.	
PA	(MATE-) INST MATERIALS RES & ENG.	
XX		
PI	Li SFY, Su X, Kwang J, Low S, Liu W,	
XX		
DR	WPI, 2001-138176/14.	
XX		
PT	Detecting the presence of an infectious agent in animals, comprises	
PT	immobilizing viral or bacterial specific antigens on a piezoelectric	
PT	crystal and measuring the change in resonance frequency of the crystal	

Haemophilus influe
High molecular wei
Non-typable Haemo
Non-typable Haemo
Haemophilus adhe
Haemophilus influe
Haemophilus influe
Haemophilus influe
Adenylate cyclase
Anti-mesothelin sc
Antimesothelin ant
Haemophilus influe
Haemophilus influe
H. influenzae BASB
H. influenzae BASB
Haemophilus influe
Haemophilus influe
Haemophilus influe
Haemophilus influe
Haemophilus influe
Haemophilus influe
High molecular wei
Drosophila melanog
Streptococcus pneu
S. pneumoniae SP06
S. pneumoniae type
Lactobacillus rham
Haemophilus influe
Haemophilus influe
Monoclonal antibod
Mouse monoclonal a
Monoclonal antibod
Haemophilus adhe
Novel human connec

Db 1 TAAKALRGAIKGKESDTPDYKGVTFITDKTGDDGNGKVTITNGEKYTLTVADIATG 60
 QY 65 ATDVNAATLQSSKNVYTSVNGQPTF 90
 Db 61 ATDVNAATLQSSKNVYTSVNGQPTF 86

RESULT 4
 AAB31434
 ID AAB31434 standard; peptide; 69 AA.

AC AAB31434;

DT 20-APR-2001 (first entry)

DE Antigenic peptide derived from the flagellin protein.

KW C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
 KM flagellin protein; poultry.

OS Salmonella enteritidis.

PN WO200078995-A1.

PD 28-DEC-2000.

PF 22-JUN-1999; 99WO-SG00061.

PR 22-JUN-1999; 99WO-SG00061.

PA (MOLE-) INST MOLECULAR AGROBIOLOGY.

PI Kwang H, Liu W, Low SS, Loh KYH;

DR WPI; 2001-071400/08.

PT New method for the specific detection of Salmonella enteritidis
 PT infections of poultry comprises contacting a biological sample with
 PT antigenic fragments of S. enteritidis fimbrial and/or flagellin
 PT proteins -

PS Claim 19; Page 43; 49pp; English.

CC AAB31433-37 represent antigenic peptides derived from the flagellin
 CC protein of Salmonella enteritidis. The specification describes a method
 CC for detecting S. enteritidis in a biological sample obtained from
 CC poultry. The method comprises contacting the sample with an antigenic
 CC fragment of S. enteritidis fimbrial or flagellin protein and detecting
 CC the formation of a complex, where the fragment is specifically recognized
 CC by S. enteritidis antibodies. The antigenic fragments are specific
 CC to Salmonella enteritidis and enable specific detection of S. enteritidis
 CC even in the presence of other Salmonella spp. The methods are useful
 CC for the specific detection of S. enteritidis infections in biological
 CC samples derived from poultry.

SQ Sequence 69 AA;

Query Match 77.4%; Score 349; DB 22; Length 69;
 Best Local Similarity 100.0%; Pred. No. 2.1e-28;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KEGDTEGYKGVTFITDKTGDDGNGKVTITNGEKYTLTVADIATGATDVNAATLQSSKN 78
 Db 1 KEGDTEGYKGVTFITDKTGDDGNGKVTITNGEKYTLTVADIATGATDVNAATLQSSKN 60

QY 79 VYTSVNGQ 87
 Db 61 VYTSVNGQ 69

RESULT 5
 AAB31435
 ID AAB31435 standard; peptide; 40 AA.

XX AAB31435;
 AC
 XX
 DT 20-APR-2001 (first entry)
 XX

DE Antigenic peptide derived from the flagellin protein.

KW C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
 KM flagellin protein; poultry.

OS Salmonella enteritidis.

PN WO200078995-A1.

PD 28-DEC-2000.

PF 22-JUN-1999; 99WO-SG00061.

PR 22-JUN-1999; 99WO-SG00061.

PA (MOLE-) INST MOLECULAR AGROBIOLOGY.

PI Kwang H, Liu W, Low SS, Loh KYH;

DR WPI; 2001-071400/08.

PT New method for the specific detection of Salmonella enteritidis
 PT infections of poultry comprises contacting a biological sample with
 PT antigenic fragments of S. enteritidis fimbrial and/or flagellin
 PT proteins -

PS Claim 20; Page 43; 49pp; English.

CC AAB31433-37 represent antigenic peptides derived from the flagellin
 CC protein of Salmonella enteritidis. The specification describes a method
 CC for detecting S. enteritidis in a biological sample obtained from
 CC poultry. The method comprises contacting the sample with an antigenic
 CC fragment of S. enteritidis fimbrial or flagellin protein and detecting
 CC the formation of a complex, where the fragment is specifically recognized
 CC by S. enteritidis antibodies. The antigenic fragments are specific
 CC to Salmonella enteritidis and enable specific detection of S. enteritidis
 CC even in the presence of other Salmonella spp. The methods are useful
 CC for the specific detection of S. enteritidis infections in biological
 CC samples derived from poultry.

SQ Sequence 40 AA;

Query Match 43.7%; Score 197; DB 22; Length 40;
 Best Local Similarity 100.0%; Pred. No. 3.6e-13;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KTGGDNGKVTITNGEKYTLTVADIATGATDVNAATLQSSKN 75
 Db 1 KTGGDNGKVTITNGEKYTLTVADIATGATDVNAATLQSSKN 40

RESULT 6
 AAB31436
 ID AAB31436 standard; peptide; 27 AA.

AC AAB31436;

DT 20-APR-2001 (first entry)

DE Antigenic peptide derived from the flagellin protein.

KW C128 fragment; fimbrial Sef14 antigen; fimbrial protein;
 KM flagellin protein; poultry.

OS Salmonella enteritidis.

PN WO200078995-A1.


```

PD 28-DEC-2000.
XX
XX 22-JUN-1999; 99WO-SG00061.
XX
XX 22-JUN-1999; 99WO-SG00061.
XX
XX (MOLE-) INST MOLECULAR AGROBIOLOGY.
XX
XX Kwang H, Liu W, Low SS, Loh KYH;
XX
XX WPI; 2001-071400/08.
XX
XX New method for the specific detection of Salmonella enteritidis
XX infections of poultry comprises contacting a biological sample with
XX antigenic fragments of S. enteritidis fimbrial and/or flagellin
XX proteins.
XX
XX Claim 21; Page 43; 49pp; English.
XX
XX AAB01433-37 represent antigenic peptides derived from the flagellin
XX protein of Salmonella enteritidis. The specification describes a method
XX for detecting S. enteritidis in a biological sample obtained from
XX poultry. The method comprises contacting the sample with an antigenic
XX fragment of S. enteritidis fimbrial or flagellin protein and detecting
XX the formation of a complex, where the fragment is specifically recognized
XX by S. enteritidis antibodies. The antigenic fragments are specific
XX to Salmonella enteritidis and enable specific detection of S. enteritidis
XX even in the presence of other Salmonella spp. The methods are useful
XX for the specific detection of S. enteritidis infections in biological
XX samples derived from poultry.
XX
XX Sequence 27 AA;
SQ
Query Match 29.5%; Score 133; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 DGNKGVSTTNGEKVTLTVADIAATGAT 66
DB 1 DGNKGVSTTNGEKVTLTVADIAATGAT 27
RESULT 7
AAB01830
ID AAB01830 standard; Protein; 1222 AA.
XX
XX AAB01830;
AC
XX
XX 11-SEP-2000 (first entry)
DT
XX
XX H. influenzae strain K1 mature full-length HMW1A protein, SEQ ID NO:37.
DE
XX
XX Mature HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;
XX non-typable Haemophilus influenzae; NTHI; non-encapsulated;
XX recombinant production; Escherichia coli; antibacterial; vaccine;
XX human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;
XX detection; diagnosis.
XX
XX Haemophilus influenzae strain K1.
OS
XX
XX Key Location/Qualifiers
XX Misc-difference 307 /note= "Encoded by CG"
XX
XX WO200020609-A2.
XX
XX 13-APR-2000.
XX
XX 07-OCT-1999; 99WO-CA00938.
XX
XX 07-OCT-1998; 98US-0167568.
XX
XX 08-DEC-1998; 98US-0206942.
XX

```

```

PA (CONN-) CONNAUGHT LAB LTD.
XX
XX Loosemore SM, Yang Y, Klein MH;
XX
XX WPI; 2000-303789/26.
XX
XX N-PSDB; AAA52180.
XX
XX Nucleic acid molecule for producing recombinant high molecular weight
XX proteins of Haemophilus which are used as a vaccine to provide
XX protection against Haemophilus induced diseases in humans.
XX
XX Claim 8; Fig 20A-R; 307pp; English.
XX
XX The invention relates to the recombinant production of Haemophilus
XX influenzae high molecular weight (HMW) proteins in Escherichia coli. The
XX expression construct used to effect recombinant expression comprises a
XX promoter functional in E. coli (e.g., the T7 promoter) operably linked
XX to a modified hmwABC operon from a non-typable (non-encapsulated) H.
XX influenzae (NTHI). Most HMW-expressing NTHI strains contain two hmw gene
XX clusters termed hmw1ABC and hmw2ABC. Each hmwABC operon comprises hmwA,
XX hmwB and hmwC genes. The hmwA genes encode the structural HMWA proteins
XX and the hmwB and hmwC genes encode accessory proteins which are
XX responsible for post-translational processing and secretion of the HMWA
XX proteins. The modified hmwABC operon used in the expression construct of
XX the invention contains an A gene modified such that it encodes only the
XX mature HMWA. The invention also discloses hmwA genes (AA52175-A52198)
XX and HMWA proteins (AAB01824-B01849) from the non-typable H. influenzae
XX strains J05c, K1, K21, LCDC2, PMH1, 15 and 12. The nucleic acids and
XX vectors are used for the production of recombinant H. influenzae HMW
XX proteins which can be used as vaccines to mediate a humoral or
XX cell-mediated immune response to provide protection against diseases in
XX humans caused by H. influenzae (e.g., otitis media, epiglottitis,
XX pneumonia and tracheobronchitis). The HMW proteins are also useful as
XX antigens in immunoassays for detecting antibodies against Haemophilus,
XX HMW proteins and/or HMW peptides. The nucleotide sequences encoding the
XX HMW proteins can be used to isolate and clone hmw genes from other
XX non-typable strains of Haemophilus via hybridisation reactions. The
XX present sequence represents a mature HMWA protein from a non-typable
XX strain of H. influenzae.
XX
XX Sequence 1222 AA;
SQ
Query Match 20.3%; Score 91.5; DB 21; Length 1222;
Best Local Similarity 24.3%; Pred. No. 1.4;
Matches 35; Conservative 15; Mismatches 37; Indels 57; Gaps 5;
QY 1 STAGTAEKATAGAIKSGKE-----GDTFDYKGV----- 30
DB 717 ATTSVEVTAKTGDIKGIISGNSGVNITAGDITLVNITGQNTVAAAGAVTTKGS 776
QY 31 -----FTIDTGTG-----DGNKGVSTT-----INSEKTLTV--ADLAT 63
DB 777 TINATGNNANITTTKGTGNGEYKASGNNITAGSNTLVNITGQNTVANSAGATTTT 836
QY 64 GATDVNAATLQSKKVVYVNGQ 87
DB 837 EGSTINATGNNANITTTGTGNGK 860
RESULT 8
AAB01828
ID AAB01828 standard; Protein; 1228 AA.
XX
XX AAB01828;
AC
XX
XX 11-SEP-2000 (first entry)
DT
XX
XX Haemophilus influenzae strain K1 full-length HMW1A protein, SEQ ID NO:34.
DE
XX
XX HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;
XX non-typable Haemophilus influenzae; NTHI; non-encapsulated;
XX recombinant production; Escherichia coli; antibacterial; vaccine;
XX human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;
XX

```


Best Local Similarity 24.1%; Pred. No. 2;
Matches 35; Conservative 15; Mismatches 37; Indels 58; Gaps 5;

QY 1 STAGTAEAKAIAAGAKGKE-----GDTEPYKVT-----30
DB 715 ATIGSEVETAKTGIDGIGESNGVNTASGDTLNVSNITQNTVVAASGAVTTKGS 774
QY 31 -----FTIDTKTGD-----DNGKYSTT-----INCKVTLTV---ADIA 62
DB 775 TINATTGNANITTTKGEINGEVKASGNVNITASGNTLNVSNITQNTVVAASGAVTT 834

QY 63 TGATDVNAATLQSSKNVYTSVNGQ 87
DB 835 TEGSTINATTGDANITTTGTGNGK 859

RESULT 10
AAB01824
ID AAB01824 standard; Protein; 1227 AA.
XX
AC AAB01824;
XX
DT 11-SEP-2000 (first entry)
XX
DE Haemophilus influenzae strain Jcyc HMW1A protein, SEQ ID NO:26.
XX
KM HMW protein; hmw gene; hmwA1; hmwA2; high molecular weight;
KM non-typable Haemophilus influenzae; NTH1; non-encapsulated;
KM recombinant production; Escherichia coli; antibacterial; vaccine;
KM human disease; otitis media; epiglottitis; pneumonia; tracheobronchitis;
KM detection; diagnosis.
XX
OS Haemophilus influenzae strain Jcyc.
XX
PN WO200020609-A2.
XX
PD 13-APR-2000.
XX
PF 07-OCT-1999; 99WO-CA00938.
XX
PR 07-OCT-1998; 98US-0167568.
XX
PR 08-DEC-1998; 98US-0206942.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Loosmore SM, Yang Y, Klein ME;
XX
DR WPI: 2000-303789/26.
DR N-PSDB; AAA52175.
XX
XX
PT Nucleic acid molecule for producing recombinant high molecular weight
PT proteins of Haemophilus which are used as a vaccine to provide
PT protection against Haemophilus induced diseases in humans -
XX
PS Claim 12; Fig 18A-R; 307pp; English.
XX
XX The invention relates to the recombinant production of Haemophilus
CC influenzae high molecular weight (HMW) proteins in Escherichia coli. The
CC expression construct used to effect recombinant expression comprises a
CC promoter functional in E. coli (e.g., the T7 promoter) operably linked
CC to a modified hmwABC operon from a non-typable (non-encapsulated) H.
CC influenzae (NTH1). Most HMW-expressing NTH1 strains contain two hmw gene
CC clusters termed hmw1ABC and hmw2ABC. Each hmwABC operon comprises hmwA,
CC hmwB and hmwC genes. The hmwA gene encodes the structural HMW proteins
CC and the hmwB and hmwC genes encode accessory proteins which are
CC responsible for post-translational processing and secretion of the HMW
CC proteins. The modified hmwABC operon used in the expression construct of
CC the invention contains an A gene modified such that it encodes only the
CC mature HMW. The invention also discloses hmwA genes (AAA52175-A52198)
CC and HMW proteins (AAB01824-B01849) from the non-typable H. influenzae
CC strains Jcyc, K1, K21, ICDC2, PMH1, 15 and 12. The nucleic acids and
CC vectors are used for the production of recombinant H. influenzae HMW
CC proteins which can be used as vaccines to mediate a humoral or

CC cell-mediated immune response to provide protection against diseases in
CC humans caused by H. influenzae (e.g., otitis media, epiglottitis,
CC pneumonia and tracheobronchitis). The HMW proteins are also useful as
CC antigens in immunoassays for detecting antibodies against Haemophilus,
CC HMW proteins and/or HMW peptides. The nucleotide sequences encoding the
CC HMW proteins can be used to isolate and clone hmw genes from other
CC non-typable strains of Haemophilus via hybridisation reactions. The
CC present sequence represents an HMW protein from a non-typable strain of
CC H. influenzae.
XX
SQ Sequence 1227 AA;
XX

Query Match 20.0%; Score 90; DB 21; Length 1227;
Best Local Similarity 24.1%; Pred. No. 2;
Matches 35; Conservative 15; Mismatches 37; Indels 58; Gaps 5;

QY 1 STAGTAEAKAIAAGAKGKE-----GDTEPYKVT-----30
DB 721 ATIGSEVETAKTGIDGIGESNGVNTASGDTLNVSNITQNTVVAASGAVTTKGS 780
QY 31 -----FTIDTKTGD-----DNGKYSTT-----INCKVTLTV---ADIA 62
DB 781 TINATTGNANITTTKGEINGEVKASGNVNITASGNTLNVSNITQNTVVAASGAVTT 840

QY 63 TGATDVNAATLQSSKNVYTSVNGQ 87
DB 841 TEGSTINATTGDANITTTGTGNGK 865

RESULT 11
AAR41731
ID AAR41731 standard; Protein; 1338 AA.
XX
AC AAR41731;
XX
DT 25-MAR-2003 (updated)
DT 26-APR-1994 (first entry)
XX
DE High molecular weight protein 3 (HMW3).
XX
KM HMW; high molecular weight protein; virus; vaccine; influenza;
KM epitope; immunity; haemophilus influenzae.
XX
OS Haemophilus influenzae.
XX
PN WO9319090-A1.
XX
PD 30-SEP-1993.
XX
PF 16-MAR-1993; 93WO-US02166.
XX
PR 16-MAR-1992; 92GB-0005704.
XX
PA (BARE/) BARENKAMP S J.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Barenkamp SJ;
XX
DR WPI: 1993-320683/40.
DR N-PSDB; AAQ49510.
XX
XX High molecular weight surface proteins - of non-typable
PT haemophilus which exhibit immunogenic properties
XX
PS Claim 5; Figure 10; 100pp; English.
XX
XX The isolation and purification of the high molecular weight protein
CC enables the identification of the major protective epitopes of the
CC protein by conventional epitope mapping. These epitopes can then be
CC synthesised using standard techniques and incorporated into fully
CC synthetic or recombinant vaccines.
CC (Updated on 25-MAR-2003 to correct PW field.)
XX

SQ Sequence 1338 AA;

Query Match 19.2%; Score 86.5; DB 14; Length 1338;
 Best Local Similarity 25.4%; Pred. No. 5.2;
 Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAFAKAIAGAIKSGKE-----GDTFDYKGV----- 30
 DB 936 ATSGTWNISRTKGTGDIKGIESTSGNVNITASGNITLKVSNITGQDVTVTADAGALTTTAA 995

QY 31 -----FTIDRTGDDGNGKYSTTINGEKVTLTVADIATGATDVNAATLQSSKNVYTS 82
 DB 996 TISATGNANITTKTGDI-NGKVESSSGSVTLVATGATLAVGNISGNTVTITADSGKLTIS 1054

QY 83 VV 84
 DB 1055 TV 1056

RESULT 12

AAAR41732
 ID AAR41732 standard; Protein; 1529 AA.

AC AAR41732;
 XX
 AC AAR41732;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-APR-1994 (first entry)
 XX
 DE High molecular weight protein 4 (HMW4).
 XX
 KW HMW; high molecular weight protein; virus; vaccine; influenza;
 KW epitope; immunity; haemophilus influenzae.
 XX
 OS Haemophilus influenzae.
 XX

PN MO9319090-A1.

PD 30-SEP-1993.

PF 16-MAR-1993; 93WC-US02166.

PR 16-MAR-1992; 92GB-0005704.

PA (BARE/) BARENKAMP S J.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX

PI Barenkamp SJ;

PT WPI; 1993-320683/40.

DR N-PSDB; AAQ49511.

PT High molecular weight surface proteins - of non-typeable
 haemophilus which exhibit immunogenic properties

XX Claim 6; Figure 10; 100pp; English.

CC The isolation and purification of the high molecular weight protein
 CC enables the identification of the major protective epitopes of the
 CC protein by conventional epitope mapping. These epitopes can then be
 CC synthesised using standard techniques and incorporated into fully
 CC synthetic or recombinant vaccines.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1529 AA;

Query Match 19.2%; Score 86.5; DB 14; Length 1529;
 Best Local Similarity 25.4%; Pred. No. 6.1;
 Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAFAKAIAGAIKSGKE-----GDTFDYKGV----- 30
 DB 1127 ATSGTWNISRTKGTGDIKGIESTSGNVNITASGNITLKVSNITGQDVTVTADAGALTTTAA 1186

QY 31 -----FTIDRTGDDGNGKYSTTINGEKVTLTVADIATGATDVNAATLQSSKNVYTS 82

DB 1187 TISATGNANITTKTGDI-NGKVESSSGSVTLVATGATLAVGNISGNTVTITADSGKLTIS 1245

QY 83 VV 84
 DB 1246 TV 1247

RESULT 13

AAAM30291
 ID AAM30291 standard; Protein; 1598 AA.

AC AAM30291;

DT 14-APR-1998 (first entry)

DE Non-typeable Haemophilus high mol.wt. surface protein HMW3.

KW Non-typeable Haemophilus; high molecular weight surface protein;
 KW HMW3; immunogen; vaccine; otitis media.

XX Haemophilus influenzae strain 5.
 XX

PH Key Location/Qualifiers

FT Misc-difference 113 /note= "encoded by GTC"

FT Misc-difference 864 /note= "encoded by TGT"

PN MO9736914-A1.

PD 09-OCT-1997.

PF 01-APR-1997; 97WC-US04707.

PR 01-APR-1996; 96US-0617697.

PA (BARE/) BARENKAMP S J.

PI Barenkamp SJ;

DR WPI; 1997-503038/46.

DR N-PSDB; AAT90992.

PT High molecular weight proteins of non-typeable Haemophilus
 influenzae - useful for vaccine production

XX Claim 1; Page 93-97; 183pp; English.

CC This protein comprises the high molecular weight surface protein
 CC HMW3 (125 KDa) of non-typeable Haemophilus influenzae strain 5 that
 CC has the immunological ability to protect against disease caused by
 CC a non-typeable Haemophilus strain and is characterised by at least
 CC one surface-exposed B-cell epitope that is recognised by monoclonal
 CC antibody AD6. The HMW3 amino acid sequence was deduced from an
 CC isolated hmw3 gene (see AAT90992). HMW1 (see AAM30293), HMW2 (see
 CC AAM30294) and HMW4 (see AAM30292) have also been identified. A
 CC conjugate comprising HMW3 linked to an antigen, hapten or
 CC polysaccharide, and a synthetic peptide of 6-150 amino acids
 CC corresponding to at least protective epitope of HMW3 are also
 CC claimed. HMW proteins, conjugates and peptides can be used in
 CC vaccines, as immunogens for preparation of antibodies and as
 CC antigens for detection of these antibodies.

XX Sequence 1598 AA;

Query Match 19.2%; Score 86.5; DB 18; Length 1598;
 Best Local Similarity 25.4%; Pred. No. 6.4;
 Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAFAKAIAGAIKSGKE-----GDTFDYKGV----- 30
 DB 1127 ATSGTWNISRTKGTGDIKGIESTSGNVNITASGNITLKVSNITGQDVTVTADAGALTTTAA 1186

Db 1196 ATSGTWNISTKTDIKIGIESTSGNVNITASGNTLKVSNIGQDVTVTADAGALTTAGS 1255
 QY 31 -----FTIDTKTGGDNGKVSSTTINGEKVTLTVADIAATGATDVAAATLQSSKNVYTS 82
 Db 1256 TISATGTNANITTKTGD-I-NKGVSSSGSVTLVATGATLAVGNISGNTVTITTAOSGKLT 1314

QY 83 VV 84
 Db 1315 TV 1316

RESULT 14
 AAW30292
 ID AAW30292 standard; Protein; 1601 AA.

AC AAW30292;

DT 14-APR-1998 (first entry)

DE Non-typeable Haemophilus high mol.wt. surface protein HMW4.

KW Non-typeable Haemophilus; high molecular weight surface protein; HMW4; immunogen; vaccine; otitis media.

OS Haemophilus influenzae strain 5.

FT Key Location/Qualifiers

FT Misc-difference 372 /note= "encoded by TCT"

FT Misc-difference 400 /note= "encoded by AAT"

FT MO9736914-A1.

PD 09-OCT-1997.

PF 01-APR-1997; 97WO-US04707.

PR 01-APR-1996; 96US-0617697.

PA (BARE/) BARENKAMP S J.

PI Barenkamp SJ;

DR WPI; 1997-503038/46.

DR N-PSDB; AAT90993.

PT High molecular weight proteins of non-typeable Haemophilus influenzae - useful for vaccine production

PS Claim 1; Page 97-102; 183pp; English.

CC This protein comprises the high molecular weight surface protein HMW4 (123 kDa) of non-typeable Haemophilus influenzae strain 5 that has the immunological ability to protect against disease caused by a non-typeable Haemophilus strain and is characterised by at least one surface-exposed B-cell epitope that is recognised by monoclonal antibody AD6. The HMW4 amino acid sequence was deduced from an isolated hmw4 gene (see AAT90993). HMW1 (see AAW30293), HMW2 (see AAW30294) and HMW3 (see AAW30291) have also been identified. A conjugate comprising HMW4 linked to an antigen, hapten or polysaccharide, and a synthetic peptide of 6-150 amino acids corresponding to at least protective epitope of HMW4 are also claimed. HMW proteins, conjugates and peptides can be used in vaccines, as immunogens for preparation of antibodies and as antigens for detection of these antibodies.

CC Sequence 1601 AA;

Query Match 19.2%; Score 86.5; DB 18; Length 1601;
 Best Local Similarity 25.4%; Pred. No. 6.5;
 Matches 31; Conservative 10; Mismatches 42; Indels 39; Gaps 3;

QY 1 STAGTAERAKA1AGAIKCKE-----GDFDYKGYT----- 30
 Db 1199 ATSGTWNISTKTDIKIGIESTSGNVNITASGNTLKVSNIGQDVTVTADAGALTTAGS 1258
 QY 31 -----FTIDTKTGGDNGKVSSTTINGEKVTLTVADIAATGATDVAAATLQSSKNVYTS 82
 Db 1259 TISATGTNANITTKTGD-I-NKGVSSSGSVTLVATGATLAVGNISGNTVTITTAOSGKLT 1317

QY 83 VV 84
 Db 1318 TV 1319

RESULT 15
 AAR9393
 ID AAR9393 standard; Protein; 2353 AA.

AC AAR9393;

DT 15-JAN-1997 (first entry)

DE Haemophilus adhesion protein HA2.

KW Haemophilus adhesion protein; HA2; hsf protein; vaccine.

OS Haemophilus influenzae type b strain C54.

FT MO9630519-A1.

PD 03-OCT-1996.

PF 22-MAR-1996; 96WO-US04031.

PR 24-MAR-1995; 95US-0409995.

PA (VYSL-) UNIV ST LOUIS.

PI (UNIW) UNIV WASHINGTON.

DR WPI; 1996-455364/45.

DR N-PSDB; AAT41476.

PT Recombinant Haemophilus adhesion proteins HA1 and HA2 - for use in vaccines against H. influenzae infection.

PS Claim 5; Page 66-73; 120pp; English.

CC Haemophilus adhesion protein HA2 (AAR9393) is associated with the formation of surface fibrils involved in adhesion to various host cells; it is also referred to hsf (Haemophilus surface fibrils).
 CC Its amino acid sequence was deduced from a genomic DNA clone (AAT41476) derived from Haemophilus influenzae type b strain C65.
 CC Large quantities of recombinant HA2 can be produced in transformed prokaryotic or eukaryotic host cells, for use in vaccines against H. influenzae infection.

CC Sequence 2353 AA;

Query Match 19.0%; Score 85.5; DB 17; Length 2353;
 Best Local Similarity 39.1%; Pred. No. 13;
 Matches 25; Conservative 6; Mismatches 28; Indels 5; Gaps 2;

QY 3 AGTAERAKA1AGAIKCKE-----GDFDYKGYT----- 30
 Db 1436 SATVSDKLSIGT-NKGVNITSTPTKGINFAKDSKTGDAN-----IHANGASTITDTULN 1490

QY 63 TGAT 66
 Db 1491 SGAT 1494

Search completed: January 12, 2004, 11:17:17

Mon Jan 12 13:45:11 2004

us-10-019-676-2.rag

Page 9

Job time : 48 secs
